DOCUMENT RESUME

ED 343 513 HE 025 380

TITLE Accreditation, Assessment, and Institutional Effectiveness: Resource Papers for the COPA Task Force on Institutional Effectiveness.
INSTITUTION Council on Postsecondary Accreditation, Washington, D.C.
SPONS AGENCY Fund for the Improvement of Postsecondary Education (ED), Washington, DC.
PUB DATE Jan 92
NOTE 75p.; For a related document, see HE 025 379.
AVAILABLE FROM Council on Postsecondary Accreditation, One Dupont Circle, N.W., Suite 305, Washington, DC 20036 ($10.00).
PUB TYPE Reports - Descriptive (141) -- Guides - Non-Classroom Use (055)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Accreditation (Institutions); *College Outcomes Assessment; Higher Education; *Institutional Evaluation; *Institutional Research; *Outcomes of Education; Postsecondary Education; Questionnaires; Research Methodology; *School Effectiveness

ABSTRACT

This resource document presents four papers that were prepared by members of the Task Force on Institutional Effectiveness of the Council on Postsecondary Accreditation concerning issues related to assessments of institutional effectiveness and the use of assessment in the process of institutional accreditation. The first paper, "Outcomes Assessment, Institutional Effectiveness, and Accreditation: A Conceptual Exploration" (Peter T. Ewell), deals with some of the basic principles and issues involved in the use of outcomes assessment in accreditation, explores some of the tensions and alternatives these create, and develops recommendations for resolution. The second paper, "Incorporating Assessment into the Practice of Accreditation: A Preliminary Report" (Ralph Wolff), reports the results of a survey of those responsible for assessment activities including their practices, their level of knowledge of assessment issues, and their perception of the effectiveness of their current outcomes assessment practices. The third paper, "Methods for Outcomes Assessment Related to Institutional Accreditation" (Thomas Hogan), reviews the various methods of outcomes assessment and discusses their advantages and limitations. The last paper, (Trudy W. Banta), represents a selected 62-item bibliography of articles and documents about outcomes assessment in higher education. (GLR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED343513

ACCREDITATION, ASSESSMENT AND INSTITUTIONAL EFFECTIVENESS

Resource Papers for the COPA Task Force on Institutional Effectiveness

★★★★★

Council on Postsecondary Accreditation
Project on "Accreditation for Educational Effectiveness: Assessment Tools for Improvement"

Partially funded by a grant from the Fund for the Improvement of Postsecondary Education.

January 1992

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY
R. Vranesh
COPA
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

HE 025 380

BEST COPY AVAILABLE

ACCREDITATION, ASSESSMENT AND INSTITUTIONAL EFFECTIVENESS

Resource Papers for the COPA Task Force on Institutional Effectiveness

Peter T. Ewell, Senior Associate, National Center for Higher Education Management Systems

Ralph A. Wolff, Associate Executive Director, Accrediting Commission for Senior Colleges and Universities, Western Association of Schools and Colleges

Thomas P. Hogan, Dean of the Graduate School, University of Scranton

Trudy W. Banta, Director, The Center for Assessment Research and Development, University of Tennessee-Knoxville

Council on Postsecondary Accreditation
Project on "Accrediation for Educational Effectiveness:
Assessment Tools for Improvement"

Partially funded by a grant from the Fund for the Improvement of Postsecondary Education

January 1992

Introduction

The four papers included in this resource document were prepared by members of the Task Force on Institutional Effectiveness of the Council on Postsecondary Accreditation. The Task Force is one of three Task Forces integral to the Council's Project on "Accreditation for Institutional Effectiveness: Assessment Tools for Improvement," a project funded in part by the Fund for the Improvement of Postsecondary Education. The papers were prepared as background papers for the work of the Task Force. However, they are also relevant to the Project as a whole in that they provide background and suggestions for further study on issues related to assessment generally and the use of assessment in the process of institutional accreditation in particular.

At the first meeting of the Task Force in March of 1991 the chairperson, James Rogers, Executive Director of the Commission on Colleges of the Southern Association of Schools and Colleges, identified four areas growing out of the charge to the Task Force which required in depth study as background for the work of the Task Force. The first of these areas involved exploring and conceptualizing more clearly the relation of outcomes assessment and analysis to institutional effectiveness and improvement in the context of accreditation. Peter T. Ewell of the National Center for Higher Education Management Systems agreed to develop a background paper on this topic and the result is the paper "Outcomes Assessment, Institutional Effectiveness, and Accreditation: A Conceptual Exploration." In the paper he deals not only with some of the basic principles and issues involved in use of outcomes assessment in accreditation but also explores some of the tensions and alternatives these create and develops recommendations for resolution.

The second area involved consideration of the effectiveness in practice of assessment and assessment related standards by institutional accrediting bodies. Ralph Wolff of the Accrediting Commission for Senior Colleges and Universities of the Western Association of Schools and Colleges undertook this assignment. As the basis for his evaluation, he developed a questionnaire which was sent to COPA recognized institutional accrediting bodies to gather information on their practices, the the level of knowledge of assessment issues of those primarily responsible for assessment activities, and their perception of the effectiveness of their current outcomes assessment practices. Dr. Wolff reports the results and their implications in "Incorporating Assessment Into the Practice of Accreditation: A Preliminary Report".

The need to develop some common understandings concerning methods appropriate to institutional assessment led to the identification of the third area for additional study. The result is the paper, "Methods for Outcomes Assessment Related to Institutional Accreditation", by Thomas Hogan, Dean of the Graduate School at the University of Scranton. In his paper, Dr. Hogan reviews the various methods of outcomes assessment and discusses their advantages and limitations. This paper will be of major value to institutions and programs as well as accrediting bodies.

Finally, the Task Force concluded that a selective basic annotated bibliography of books, articles and documents would be of great value to accrediting bodies and institutions. Trudy W. Banta, Director of the Center for Assessment Research and Development of the University of

Tennessee - Knoxville, agreed to undertake this assignment. The bibliography is designed to provide those references that will give accrediting agency staff and those with whom they consult a comprehensive overview of the current status of the field of postsecondary student outcomes assessment and the assessment of institutional effectiveness. The paper produced by Dr. Banta "Selected References on Outcomes Assessment in Higher Education: An Annotated Bibliography" serves that function extraordinarily well.

These papers greatly facilitated the discussions of the Task Force and the preparation of the "Report of the Task Force on Institutional Effectiveness." We believe that they also constitute a major contribution to the literature on assessment and that they will be of direct help to accrediting bodies and institutions in reviewing or developing their own assessment activities. On behalf of the Council on Postsecondary Accreditation, the Project and the Task Force, we express our thanks and appreciation to Drs. Ewell, Wolff, Hogan and Banta for their work in preparing these papers.

Marianne R. Phelps
Senior Associate
Project Director

Richard M. Millard
Project Coordinator

Table of Contents

Introduction **00**

Outcomes Assessment, Institutional Effectiveness, and Accreditation: A Conceptual Exploration, Peter T. Ewell **1**

- Some Key Concepts and Their Relations 1
  - A Possible Framework for Discussion 2
    - *Figure 1*: A Conceptual Typology of Terms Relating to Institutional Outcomes, Effectiveness, and Accreditation 2
  - Outcomes Assessment as a Concept 4
    - Individual or Institutional Unit of Analysis 4
    - Indicators or Standards 4
    - Attained Outcome or Student Growth and Development 5
    - Making Judgments or Making Improvements 5
    - Institutional or Constituent Perspectives 6
    - Long-Term or Short-Term Perspectives 6
  - Institutional Effectiveness as a Concept 7
    - General Goal Attainment or Specific Institutional Attributes 7
    - Any Mission or a Specifically Educational Mission 7
    - Any Outcomes or Specific (Common) Outcomes 8
  - Institutional Accreditation as a Concept 9
    - Validating "Quality" or Validating the Quality-Assurance Process 9
    - Certifying Quality or Stimulating Improvement 10
    - Source of Legitimacy: Academic Community or Society and the Wider Public 10
    - Focus of Concern: The Institutional Community as a Whole or the Sum of the Disciplines and Subunits that Comprise it 10
    - *Figure 2*: An Evolving Role for Accreditation? 11
- Some Operational Dilemmas 12
  - How to Avoid Assessing Everything 12
  - How to Arrive at Reasonable Expectations Regarding Outcomes Assessment in the Process Itself 13
  - How to Best Communicate the Results of Assessment 13
  - How to Get Institutions to Take the Process Seriously 14
  - How to Sustain Institutional Action and Attention to Assessment Over Time . 15
  - How to Link Accreditation's Role in Outcomes Assessment with Those of Other Processes 15
- Some Recommended Principles for Future Policy 16

Incorporating Assessment into the Practice of Accreditation: A Preliminary Report, Ralph A. Wolff **19**

- Design of the Survey 19

Survey Results 20
  Respondent's Personal Knowledge and Experience with Assessment 20
  Knowledge and Experience of Association Staff and Commission Members . . 20
  Standards and Guidelines on Assessment 21
    Team Training Workshops 21
    Self Studies 22
    Team Visits 22
    Commission Actions 22
  Personal Views of Respondents 23
    Additional Support Desired 24
Analysis and Conclusions 24
Tabulated Responses 26

Methods for Outcomes Assessment Related to Institutional Accreditation, Thomas P. Hogan 37

Preface 37
Introduction 37
  The First Key Question: Final Status or Value Added 39
  The Second Key Question: Final Answers or Meaningful Processes 39
  Program Effectiveness vs. Institutional Effectiveness 41
  The Importance of Goals 41
  A Note on Reliability and Validity 42
Methods for Assessing Outcomes 42
  Tests of Developed Abilities, Knowledge, and Skills 43
    Areas to be Covered 43
    Sources of Development 43
    Externally Developed Tests of General Education 44
    Locally Developed Tests of General Education 45
    Externally Developed Tests of Specific Areas 45
    Locally Developed Tests of Specific Areas 45
    A Special Note on Portfolios 45
    General Strengths and Weaknesses of Tests 45
    Strengths/Advantages of Externally Developed Tests 46
    Weaknesses/Disadvantages of Externally Developed Tests 47
    Strengths and Weaknesses of Locally Developed Tests 47
  Surveys, Questionnaires 48
    Sources of Development 49
    Target Groups 49
    Format 49
    Content 50
    Strengths/Advantages of Surveys 50
    Weaknesses/Disadvantages of Surveys 52
  Institutional Records 52
    Strengths/Advantages of Institutional Records 53
    Weaknesses/Disadvantages of Institutional Records 53

        Concluding Note on Institutional Records 54
    The Institutional Assessment Plan 54

Selected References on Outcomes Assessment in Higher Education: An Annotated Bibliography, Trudy W. Banta 57
    Forward 57
    Assessment and Accreditation 57
    Measurement Issues 58
    State Policy Issues in Assessment 60
    Assessment in the Major 61
    Assessment in General Education 62
    Classroom Assessment Techniques 63
    Developing Goals for Assessment 63
    Assessment of Student Development 63
    Assessment in Community Colleges 64
    Books/Collections/Review Articles on Assessment Topics 64
    Assessment Bibliographies 67
    Assessment Periodical 67

Outcomes Assessment, Institutional Effectiveness, and Accreditation: A Conceptual Exploration

Peter T. Ewell*

In the last three years, most institutional accrediting bodies have adopted a policy on the assessment of institutional outcomes. Some, like the Southern Association of Colleges and Schools (SACS) and the Western Association of Senior Colleges (WASC), have made "institutional effectiveness" a distinct standard for accreditation. Others, including the North Central Association (NCA), the Middle States Association, and the New England Association, have made the assessment of student and institutional outcomes a part of more general standards regarding the achievement of institutional purpose. In part, these actions reflect an original and historic interest in the quality of educational results on the part of the accrediting community. In part, however, they constitute a reaction to a set of recent events—particularly, the rapid development of a national "assessment movement" and the growing insistence of government authorities (both state and federal) that colleges and universities become more accountable to the public for their educational products.

Because of this rapidly developing background, there has been little opportunity for members of the institutional accrediting community to develop from first principles a systematic conception of the proper role of assessment in the process of which they are a part. This paper represents an initial attempt to address this condition. It contains two main sections. The first presents a discussion of the relationships among key terms: "outcomes assessment," "institutional effectiveness," and "accreditation." A primary premise is that such terms currently have multiple meanings and it is difficult to move forward until these are systematically recognized and resolved. The second section presents a set of more specific operational dilemmas that typically arise whenever "assessment" and "accreditation" come together. This list is based on the emerging experience of both institutions and associations, and each is briefly discussed with respect to origins, specific manifestations, and suggested remedies. A final section proposes some initial recommendations for action on the part of institutional accrediting bodies, based on both prior discussions.

Some Key Concepts and Their Relations

A challenge facing any commentator on assessment is one of definition. Through popular usage, the term has come to mean so many different things in different contexts that clear communication is threatened. The same is true of such terms as "quality" and "institutional effectiveness." A first step in determining the proper role of assessment in institutional accreditation, therefore, is to explore more specifically exactly what we are talking about.

Two common difficulties here are distinguishing more clearly between results and processes, and better specifying the intended unit of analysis. "Institutional effectiveness," for instance, refers to a particular *property* of an *institution*, while "institutional assessment" refers to the *means* by which we determine that this particular property is present. Yet the primary center of gravity of many current discussions of "institutional effectiveness" in accreditation circles (particularly, of course, in the southeast) has been much more narrowly focused: what measures or procedures will provide acceptable evidence of student learning? Posing the problem in this way both shifts the concept's domain and its original unit of analysis. At the same time, the key terms in this relation themselves need further conceptual clarification. Given this challenge, this section first describes some of the overall conceptual difficulties associated with relating the three terms "assessment", "institutional effectiveness", and "accreditation"; it then goes on to explore some particular difficulties associated with each component concept.

1. ***A Possible Framework for Discussion***. Figure 1 presents a first attempt to sort through some of this terminology. On one dimension, it distinguishes the particular *properties* of interest ("domains") from the *process* used to determine the degree to which such properties are present. Properties are chiefly results or conditions—attributes that apply to institutions, programs, or individuals. "Processes" consist of various forms of evidence-gathering used to ascertain or to certify that these results or conditions exist. The diagram's second dimension emphasizes the fact that this distinction can operate simultaneously at many levels and that considerable conceptual confusion can arise from the fact that similar terms are used to describe quite different things.

Figure 1
A Conceptual Typology of Terms Relating to Institutional Outcomes, Effectiveness, and Accreditation

Level of Conceptual *Specificity*:	Conceptual Domain (A *Result* or *Condition*)	Process to Determine Current *State* of Domain
1. Overall "Quality" ↓	"Accreditable Standard of Quality	Accreditation Process
2. Mission/ goal Attainment ↓	"Institutional Effectiveness	"Assessment (1)"
3. Educational (Student) Results ↓	"Student Outcomes"	"Assessment (2)"
4. *Cognitive* Results	"Student Learning Outcomes"	"Assessment (3)"

As the figure indicates, at least four such levels, arranged in a descending hierarchy of specificity, are usually present. The first refers to the "quality" of an institution as a whole—public validation of which is a primary object of accreditation. Embedded in the process of self-study and peer review, the precise content of this domain is appropriately broad and deliberately vague; typically, however, it embraces under a single conceptual umbrella such diverse aspects of a college or university as the appropriateness and clarity of its mission, the adequacy of its resources and underlying physical and organizational structures, and the results that it achieves. Mission or goal attainment, in contrast, is a far more circumscribed concept, noted explicitly as "institutional effectiveness" in SACS and WASC standards, and implied by other regional bodies through statements such as "the institution is achieving its purposes" (North Central Association Evaluative Criterion Three, 1988-90). While the unit of analysis here remains the institution as a whole, this level of specificity embraces outcomes to the virtual exclusion of resources and processes.

Moving to the third level of specificity—educational results—is critical, as it entails both further circumscription of the conceptual domain and an implied shift in the operational unit of analysis. Though they are a partial result of what institutions do, educational results actually happen to students. Investigating them will necessarily involve looking at individuals, though the purpose remains to aggregate obtained results to inform wider judgments about how well programs and institutions are functioning. Moving to the fourth and final level is then straightforward: only *cognitive* results of collegiate instruction are considered, but the operational assessment processes required are very similar.

A major current conceptual problem of assessment in the institutional accreditation process is that levels two through four are not well distinguished and it is not clear which is the primary focus of the process. Partly this is because educational results, and particularly student learning outcomes, represent the one clear common denominator among colleges and universities with otherwise vastly different missions and student clienteles. All institutions should therefore pay some attention to them in approaching accreditation. Partly it is because obtained results (and particularly student learning outcomes) have historically been neglected in accreditation, and are consequently now being particularly emphasized in reaction. The term "assessment", moreover, appears currently to be used almost interchangeably across all three levels. As noted in figure 1, it is currently applied to the process of gathering evidence about institutional goal attainment, the process of documenting student outcomes of all kinds, and the process of determining precisely what students have learned.

Given this situation, one step toward clarifying the relationship between assessment and institutional accreditation might be to better emphasize in all communications and guidelines that level two constitutes accreditation's center of attention. "Institutional effectiveness" requires active demonstration that claimed results are being achieved, and this means that *concrete evidence* of goal achievement must be visibly included in the results of self-study. "Assessment (1)" therefore requires actual evidence to be presented during self-study, while "Assessment (2)" and "Assessment (3)" can be primarily presented and examined as *processes*. Consistent usage might then reinforce the notion that the institution remains the primary unit of analysis in institutional accreditation: all "assessment" should clearly support this common purpose regardless of its operational level of data collection. Consistent usage would also help to clarify the point that while educational results are indeed important as the common products of all college and university activity, they are not the only aspects of institutional goal-attainment that should be investigated. Evidence of attainment around other goals should also, where appropriate, be provided.

According consistent conceptual primacy to level two requires institutional accrediting bodies to make two deceptively simple further statements about what should happen in the accreditation process. First, accreditation guidelines should emphasize that the goal attainment or "institutional effectiveness" constitutes only a part of what's required under an "accreditable standard of quality." Impressive outcomes may be purchased in the short term at the price of deferred attention to the maintenance of key institutional assets, imperiling long-term viability. Or goals themselves, though attained, may be inappropriate or contradictory.

Second, such statements should make clear that when it occurs in conjunction with accreditation, student outcomes assessment must clearly focus upon establishing the institution's particular *contribution* to the outcomes achieved. Simply running a testing program for a variety of operational purposes is not enough. This point speaks to a particular operational confusion that appears to be currently troubling accreditation in the realm of assessment: confounding evidence-gathering procedures established to periodically monitor and improve instruction with those intended to certify student attainment (such as exit examinations) or to accord appropriate instructional treatments to individual students (such as placement examinations). Because institutions with the latter are "practicing assessment," it is sometimes asserted that they are "investigating effectiveness" though the results are never aggregated to the program or institutional level and are never used to actively guide instructional improvement. While programs such as these are surely important, their proper place in the accreditation process is in the analysis of an institution's curriculum and instructional practices, not in an assessment of its effectiveness.

2. ***Outcomes Assessment as a Concept.*** Reflecting its evolution in higher education generally, outcomes assessment in the accreditation community is a relatively recent concern. While most institutional accreditation standards before 1985 in some way referenced the achievement of institutional purposes, few in practice required institutions to provide explicit evidence of educational goal achievement. Since that time, in parallel with assessment's growing prominence in state and federal higher education policy, virtually all have adopted "assessment" requirements. But in many cases these requirements were adopted hastily. Accrediting bodies were in part concerned that the process of voluntary accreditation not be left behind in a political world that was increasingly emphasizing concrete results as the proof of "quality." In this press to adopt standards, however, important conceptual issues inherent in the concept of assessment remained obscure. To evolve an appropriate approach for the future, the institutional accreditation community must examine each of these issues systematically, and must make appropriate collective choices about the appropriate path to take.

Six major conceptual tensions embedded in assessment appear particularly relevant. Each describes a continuum of choices about how to articulate, recognize, and enforce assessment standards in the process of self-study. In some cases, appropriate choices for accreditation seem clear, while in others, the determination of the best location on the continuum will depend upon institutional mission and context. Each issue is listed below, together with a brief discussion of its implications for self-study:

- *Individual or institutional unit of analysis?* The term "assessment" is currently being applied to widely disparate units of analysis including states, institutions, programs, and individual students. As a result, an enormous range of practices within accreditation currently occur under a common label. Assessment practices submitted as evidence of compliance with SACS "institutional effectiveness" standards and NCA "achievement of institutional purpose" standards can vary from occasional large-scale, sample-based alumni surveys, through basic skills placement examinations regularly administered as a matter of policy to designated students, to isolated "classroom research" projects designed and executed by individual faculty members and academic departments for their own purposes. All such practices are exemplary and should be appropriately recognized. But not all constitute evidence of institutional effectiveness per se. Accreditation standards and practice, therefore, should primarily emphasize processes intended to provide evidence of institutional goal attainment. This is the level at which concrete evidence should be present and examined in self-study. Individual and classroom assessment, if present, should be noted and discussed as processes and included in the self-study's description of curriculum and academic practice.
- *Indicators or standards?* Partially as a result of confusions regarding unit of analysis, the use of assessment evidence in accreditation often entails a further difficulty: are stated outcomes intended as minimum standards or instead as indicative of broad mission attainment? In the first case, the focus of assessment must be to determine that all students in fact meet the standard—for example, in the form of a given level of assessed writing competence or a knowledge-based major field exit examination. In the second case, the focus of assessment is to provide evidence that the "typical" student or graduate embodies a set of characteristics that the institution seeks to produce more broadly—though particular individuals may vary

considerably in the degrees to which these characteristics are present and the ways in which they are manifested. (The situation is similar to that of stated admissions requirements with respect to test scores. Minimum cutoffs may exist below which the institution will not accept an applicant, but the scores of admitted students may vary considerably around a stated average.) Accreditation may require both kinds of assessment, but in different places and for different purposes. Where an institution states minimum standards for learning outcomes—and there ought always to be some of these present—it has an obligation to demonstrate that it has in place an adequate mechanism to determine if outcomes that meet these standards are in fact being obtained, and must describe this mechanism as part of its educational program. For broader outcomes goals (for instance, those typically contained in a mission statement), what is needed in self-study is evidence that aggregate educational results are consistent with these goals. Evidence of this kind can appropriately be collected periodically and may show substantial variation across students; what's important for "institutional effectiveness" is what the central tendency shows.

- *Attained outcome or student growth and development?* A slightly different issue is the degree to which, through accreditation, institutions should be held accountable for certifying stated outcomes (regardless of the degree to which they may have actually helped to produce them) or for "developing talent" through their activities (regardless of the final levels of achievement reached by students at the end of their programs). This is an underlying issue in assessment itself, embodied in an ongoing debate about the appropriateness of "value-added" studies. Most discussion here occurs at the individual unit of analysis and at this level it indeed represents a legitimate analytic choice: should assessment provide evidence of whether students possess key skills at graduation or should it instead be focused upon determining the institution's contribution to the development of these skills? The answer to this question will depend entirely upon the purpose of assessment. As evidence of "institutional effectiveness," however, both are generally held to be important, and the assessment process should reflect this. As part of self-study therefore institutions should be able to present acceptable evidence both that broad goals are in fact typically being attained and that the actions of the institution had something to do with them. This need not (and in most cases ought not) imply a "test-retest" design as is sometimes advocated. But it does require some consideration when discussing outcomes of both entering student abilities and of actual educational processes. To be really adequate, in fact, it requires some demonstration of the way educational processes connect to intended outcomes, both in concept and in fact.
- *Making judgments or making improvements?* Advocates of institutional assessment in higher education increasingly emphasize its contributions to instructional improvement. Engaging in the process itself, it is argued, can result in such major benefits as clearer goals and better common understanding of instructional activities across programs and institutions, often regardless of what is found. Partly as a result, institutional assessment practice is increasingly shifting toward locally-developed instruments and approaches in which such benefits can best be realized. Accountability-oriented processes, in contrast, tend to rest upon common standards and assessment procedures —often embodied in a standardized examination. Because the process of institutional accreditation is simultaneously formative and summative, managing this dichotomy can

pose particular difficulties, and ideally, evidence of both kinds should be required. "Achievement of purpose" demands that the institution build a convincing case that results are consistent with intentions. At the same time, the process of assessment should be one in which many members of the campus community are involved and, through such engagement, should be one through which the institution can demonstrably learn.

- *Institutional or constituent perspectives?* Assessment practitioners are also increasingly being forced to recognize that judgments about goal attainment may differ considerably depending upon where one sits. What a faculty thinks is important for a given graduate may or may not correspond to what an employer thinks is necessary or, indeed, to what a student actually wants. How much legitimacy should each of these perspectives be accorded in assessment, particularly when it is conducted in the context of accreditation? Appropriate resolutions may depend largely on the content of the institution's mission and its principal indicated constituencies. Institutions whose missions visibly embody a "service" ethic—for example, community colleges or regional public institutions—should arguably take special care to demonstrate that the perspectives of students and employers have been taken into account and to what degree these "customers" are satisfied with what they have received. Entering student goal studies, information about student and alumni satisfaction, and employer needs assessment should be particularly prominent in this respect. Because no sector of higher education exists in a vacuum, moreover, all institutions ought appropriately to recognize the existence of external clients and how well they are served, be they students, other postsecondary institutions, or the public at large.
- *Long-term or short-term perspectives?* Many of the most important outcomes of higher education, it is commonly claimed, take a long time to become manifest. Determining an appropriate time-frame for assessment is therefore often an issue. Institutions may claim, for example, that they should not be held directly accountable for "life-long learning" until sufficient time has elapsed for this phrase to become meaningful. On the other hand, they may contend that it is unfair to look at graduate abilities long after graduation because a variety of additional experiences may have mitigated or modified what was once learned in college. From the standpoint of accreditation, three points with respect to this issue seem clear. First, as in so many such issues, the institution's own goal statements must be taken as a guide. Phrases such as "job-entry skills" or "lifelong learning" in themselves suggest appropriate timeframes and the institution's assessment procedures should reflect them. Secondly, the timeframe chosen should reflect the primary intent of self-study as part of a process of institutional improvement. Very little of utility for current operations can generally be learned from an assessment of students who attended the institution long ago under very different contextual conditions. As a consequence, whatever the outcome, the timeframe chosen for its assessment should demonstrably result in information that is relevant to current circumstances. Finally, because instruction is a process that occurs over time, observations about student development over time are particularly salient. As a result, assessments occurring at more than one point or as a product of longitudinal studies may be especially useful.

Choosing an appropriate location on each of these six dimensions must precede discussions of assessment methodology in the context of institutional accreditation. In the past, both institutions and accrediting agencies have tended to move too quickly to technical issues when discussing assessment, and have invested too little time in determining what should be done about these more basic conceptual issues. All should be recognized more explicitly in regional accreditation guidelines and where needed, in better directions to visiting teams, and in self-study workshops.

3. ***Institutional Effectiveness as a Concept.*** Like the concept of "assessment", the history and conceptual clarity of the term "institutional effectiveness" is murky. While the term is now explicitly recognized in the standards of at least two regional accreditors, and is implied by the guidelines of all others, it continues to mean quite different things to different people. In parallel with the notion of assessment, three particular conceptual trade-offs have a bearing on the process of institutional accreditation and must be specifically addressed:

- *General goal attainment or specific institutional attributes?* While goal attainment remains the center of gravity for accreditors when discussing the term "institutional effectiveness," it is important to remember that researchers in both higher education and in organizational theory identify at least three additional dimensions the term, all of which are in some way relevant to the accreditation process. *Managerial process* views of effectiveness stress the organization's concern with efficiency through the use of such "best practice" managerial techniques as clear organizational structures, lines of communication, efficient allocation and use of resources, and availability and use of information. Concern for such processes pervade most self-study guidelines and in the "institutional research" clause, are also present in the "institutional effectiveness" standards of SACS and WASC. *Organizational climate* views of effectiveness emphasize the supportive aspects of the institution's culture from the point of view of those who work there, and typically assess such elements as the perceived reward structure, overall employee satisfaction and morale, and loyalty to the institution. Though less cited in self-study guidelines, such considerations are often among the paramount concerns of visiting teams in the accreditation process. Finally, Environmental adaptation views of effectiveness stress the organization's ability to acquire and maintain resources, and its vulnerability to environmental change; key elements of assessment here include analyses of "market niche", of particular patterns of community and constituency support, and the institution's ability to innovate. Again, such concerns pervade accreditation, embodied, for instance, in NCA's standard four, which states "the institution can continue to accomplish its purpose." Clearly, the process of institutional accreditation is concerned with all four of these things. Guidelines that explicitly reference "effectiveness", however, are for the most part focused only on the attainment of formally established goals. While studies of goal attainment, and particularly of educational results, should remain a central concern, care should be taken by accrediting bodies that broader meanings of effectiveness are not thereby obscured. Ideally, these broader meanings should permeate both the self-study process and the document that results from it.
- *Any mission or a specifically educational mission?* Centering the concept of "effectiveness" on institutional goal attainment, however, leaves unanswered the further question of whether any mission will do. Most colleges and universities are in the business of instruction, but they devote varying amounts of attention to this assignment in their mission and goals statements. If the attainment of stated goals is taken as the sole and literal measure of institutional effectiveness, institutions will argue that they

should not be penalized for failing to provide evidence about what they consider to be low priority activities. Should assessments of effectiveness be limited solely to what an institution says is important, or should it also include an analysis of the goals themselves?

Accreditation processes provide two points of attack on this issue. First, most do include an analysis of goal statements with regard to appropriateness and clarity. Review standards that require an explicit analysis of the prominence, adequacy, and appropriateness of posed instructional goals might be especially emphasized here. Philosophically, of course, this begins to change the nature of the accreditation process from institutional validation toward standard setting—a shift which must be carefully considered. Behaviorally, however, the fact that all institutions will have some mention of instruction in their missions means that attention can be largely directed toward making these statements more visible and more explicit. Once this is accomplished, assessment follows naturally. On the other hand, institutions can be required to explicitly examine the results of instruction regardless of the content of their goal statements—in effect, the requirement of SACS standard three. Here, however, institutions may develop statements of intended instructional outcomes that are not actually present in their mission statements. While this approach has the virtue of not allowing the issue of instructional effectiveness to be ignored, the danger is that "assessment" may be seen as a separate process, unconnected to the remainder of the self-study. On balance, taking a more proactive stance regarding institutional goals and their appropriate contents seems the more appropriate for accreditation, though both approaches appear useful.

- *Any outcomes or specific (common) outcomes?* Similarly, because of their common engagement in instruction, should all colleges and universities be held accountable for the attainment of *particular kinds* of learning outcomes? With the recent development of several common state-level assessments of general education and increasing momentum for an assessment of collegiate-level "communication, critical thinking and problem-solving abilities" at the federal level, this is a far more controversial issue, and an increasingly insistent one as well. Most colleges and universities teach not only undergraduate programs, but also some form of general education. Regardless of explicit goals statements, should they be required, through accreditation, to assess the results of this activity according to a set of recognized common attributes (for instance, critical thinking or communications ability)? In parallel, some accrediting agencies—most notably WASC in the area of diversity—are requiring that institutions pay special attention to particular issues. Should this extend to the curriculum, and more explicitly to educational outcomes? In contrast to specialized or professional accreditation where certifying the attainment of specific skills may be paramount, the best course for institutional accreditation is probably to address this topic through existing goal statements. It is unlikely, for instance, that any institution for which these are appropriate will avoid any mention of general education or intended generic outcomes of college in its mission statement. Guidelines should ensure that, when present, such statements are clear and are ultimately amenable to assessment. They might also require institutions to address specific *kinds* of generic outcomes, such as oral communications or knowledge of other cultures. But they should probably stop short of prescriptions of the exact levels of knowledge and skills that ought to be present.

All three of these issues are, of course, heavily centered on the instructional function of colleges and universities—particularly at the undergraduate level. This should not, however, be taken to mean that the domain of institutional effectiveness is limited to these areas. Many additional goal areas of higher education can and should be addressed under the rubric of institutional effectiveness including research, public service, and the development of particular values consistent with the institution's mission. A good point of departure in this regard, suggested by several commentators, is to examine and analyze the institution's "implied promises", as noted in such publications at the catalogue, recruitment view-books, publicity materials, and the like. Taken together, these materials often suggest an additional set of "thematic goals" in terms of which to structure evidence-gathering for demonstrating institutional effectiveness.

4. ***Institutional Accreditation as a Concept.*** The process of institutional accreditation has also embodied conceptual tensions from its outset. At its best, accreditation is a complex, multi-faceted process in pursuit of multiple purposes. It is simultaneously the academy's primary mechanism for quality assurance and one of its major potential avenues for self-improvement. It is therefore a process that must carefully balance internal and external forces in determining what is appropriate. Finally, it is a process in which evidence is both critically important and is largely nonprescribed; in the tradition of self-study, it is equally unthinkable for the process to avoid investigating important areas of institutional functioning and for it to dictate what the contents or form of investigation should look like.

These characteristics of institutional accreditation are in marked contrast to apparently similar processes. Special or programmatic accreditation process—though they also have dual purposes—are far more directed toward the task of assuring compliance with minimal standards. Government regulation of public higher education—though equally concerned with accountability and quality improvement—is of necessity far more focused on the achievement of public purposes and the determination of return on public investment. For both these processes, evidence of results is crucial in itself. For institutional accreditation, such evidence is more valuable for the larger insights which it can provide about how the institution as a whole is functioning and how it might function better.

As institutional associations increasingly engage in outcomes assessment, many of the most basic conceptual issues inherent in the accreditation process itself reappear in new guises. Among the most compelling of these are the following:

- *Validating "quality" or validating the quality-assurance process?* The "bottom-line" of institutional accreditation is summative: a community judgment regarding the overall integrity and viability of a college or university as an educational enterprise. Yet its primary mechanism is not inspection but self-study; the role of peer review is to ensure that the self-study processes used are undertaken fairly and appropriately, and not typically to render a set of wholly independent judgments. So long as the primary domain of self-study remained resources and processes, the conceptual tension between these two elements was muted. With outcomes as a focus, however, the issue is more salient. At the very minimum, submitted evidence must be directly examined by outsiders, if only in comparison to stated goals. At the same time, the assessment processes used must be externally validated and, as most regional accreditors are currently discovering, this is often a far more complicated process from both a technical and an organizational standpoint than is typical in reviewing the results of more traditional self-study.

- *Certifying quality or stimulating improvement?* A somewhat similar issue is the question of the degree to which the accreditation process is intended to serve primarily as a means for determining quality or as a mechanism for improvement. Both are important and both are traditionally claimed as valuable results of self-study and peer review. Again, however, explicit consideration of outcomes stretches the ability of a single process to do both jobs. The assessment mechanisms that are best suited to demonstrating effectiveness are not always those that are the most helpful in the long run for program improvement. Nationally normed standardized examinations may provide greater external credibility and important comparative information, despite their common inability to provide useful guidance at the local level. Similarly, if the object is to provide a convincing demonstration of goal attainment, an institution may be better advised to invest its limited analytical resources in a couple of periodic, well-designed, methodologically-sophisticated assessment efforts than in a much wider range of less elaborate but more frequent evaluative efforts. Finally, quite different methodologies may be required to demonstrate the attainment of defined standards than to show continuous improvement. Current accreditation statements about outcomes often do not provide sufficient guidance about which course to take: "evidence of achieving institutional purpose" suggests one set of choices while "creating a culture of evidence" may suggest one that is quite different.
- *Source of legitimacy: academic community or society and the wider public?* By its very nature, accreditation must play a mediating role between society and the academic community. On the one hand, it is owned and operated by higher education and as preeminently peer review, it must largely reflect the standards and the values of the academy. On the other hand, it is a process whose message is directed outward and in which the public is traditionally asked to place its trust. As assessment mandates become pervasive, moreover, the latter role is increasingly shared with state and federal authorities. So long as resources and processes constituted the primary domain of accreditation, primary questions of efficiency and access were of interest to both parties. As results become equally paramount, however, judgments about both their appropriate content and their quality may legitimately differ. Current accreditation language for the most part leaves this choice up to the institution but given current political circumstances, this may be a position that is increasingly untenable.
- *Focus of concern: the institutional community as a whole or the sum of the disciplines and subunits that comprise it?* Institutional accreditation is also one of the few processes in the academy that can force consideration of colleges and universities as complete entities. Special or professional accreditation tends to examine the functioning of units and programs in isolation and often, it is claimed, at the expense of overall institutional goals and priorities. State oversight through coordinating and governing boards, is often similarly focused on subunit operations—particularly in its most common guise of program review and approval. Given this unique potential, should institutional accreditation be primarily concerned with examining the entire academic community and how it works? Again, the explicit introduction of assessment into the process helps make this choice explicit. Many institutions—particularly large research universities where departmental and disciplinary cultures are stronger than allegiance to the institution as a whole—present as evidence of overall "effectiveness"

largely the results of major field assessments, with little attention paid to what happens in general education or in the wider institutional environment. Current accreditation guidelines, with a few exceptions, do not address this issue explicitly: institutions are expected to produce assessment evidence at both the general and programmatic levels, following the logic of established goals. But given institutional accreditation's unique focus on the complete academic community, consideration might be given to strengthening this component of evaluation by requiring institutions to specifically attend to cross-departmental "joint products" in their assessments.

In all four areas, it appears, considering outcomes renders more explicit some conceptual issues that have always been present in the accreditation process. At the same time, considering outcomes may signal a changing role for the process itself. Figure 2 attempts to capture this evolution by contrasting the implications for appropriate evidence of two kinds of theoretical accreditation processes. The first, or "traditional" role has as its primary focus validating overall institutional quality from the standpoint of the academy. The second, or extended role adds the notion of validating the process of quality assurance operated by the institution, but from the perspective of the wider society. The underlying philosophical shift is subtle but important. From simply reflecting the values of an academic community, the process involves by implication an attempt to actively *shape* those values where they critically interact with society. At the same time, the critical underlying source of credibility for the process shifts from exclusive reliance on a set of implied but rigid standards of quality to include an additional set of pervasive institutional attitudes and activities visibly directed toward improvement.

Figure 2
An Evolving Role for Accreditation?

	Traditional Role	Extended Role
PRIMARY FOCUS OF PROCESS:	Validating "*Quality*" from the standpoint of the *Academic Community*	Validating "*Quality Assurance*" from the standpoint of *Society*
IMPLICATIONS FOR EVIDENCE:		
Content of Information:	Goals, resources and processes (in relation to "industry standards")	Resources, processes and results (in relation to goal achievement"
Use of Information:	Judgmental	Indicative
Informational Perspective:	The Academy	Constituents
Informational Focus:	Minimum Standards	Continuous Improvement

This shift has much in common with the recent "quality revolution" in business and industry, embodied in such organizational review processes as that associated with the Malcolm Baldrige National Quality Award, and it has profound implications for both the nature of evidence and its use in improvement. As noted in Figure 2, at least four such implications are relevant. First, the content of information moves from a comparison of institutional goals, resources, and processes with implied fixed standards to include a comparison of obtained results with established goals. Second, the primary role of evidence is to indicate progress rather than to certify attainment, and *patterns* of indicators are intended to be used in combination to suggest effectiveness rather than to establish piecemeal the degree to which individual standards are met. Third, provided evidence should indicate a concern with the institution's "customers" (students, employers, and community) as well as its own members. Finally, the primary case to be made is less the fact that the institution meets minimum fixed criteria than the fact that it has the capacity, the will, and the culture to continuously improve.

This evolving role for institutional accreditation is intriguing because, like industry's Baldrige Award, it is simultaneously outwardly-directed and founded on peer review. Governmental assessment mandates, though in most cases intended to stimulate improvement, appear unusually vulnerable to political pressures that threaten to turn them exclusively into accountability processes emphasizing conformance to fixed standards (the evolution of Tennessee's performance funding program provides a case in point). Because of their high stakes, it is often hard for such processes to establish an ongoing "culture of evidence," despite their best intentions. By concentrating greater effort on validating an institution's quality assurance processes from a wider community standpoint, accreditation might take on a role that is both badly needed and for which it is ideally suited. No other process in the academy has currently claimed this potential "market niche." By reason of its history, culture, and primary attributes, institutional accreditation may well find its best future in this activity.

Some Operational Dilemmas

Regardless of the degree to which the conceptual evolution described in Figure 2 constitutes a real change in kind, the introduction of outcomes assessment into institutional accreditation raises significant operational dilemmas regarding how the process should be conducted. Some result directly from the new challenges posed by assessment; some are traditional difficulties that are particularly exacerbated by the presence of assessment. All, however, appear to be difficulties that are increasingly being faced in actual practice across a wide range of institutional accrediting associations and types of institutions.

Among the most persistent such issues are the following. Each is rooted in more basic conceptual ambiguities noted earlier, but each is also an operational problem worth discussing in its own right.

1. ***How to avoid assessing everything.*** If goal achievement becomes the primary mark of institutional effectiveness and as a consequence, if assessment becomes the primary means for demonstrating quality, must an institution closely and formally examine *every* established goal and its implications? This problem is increasingly apparent in assessment generally, and is widely reported as an initial stumbling block by institutions attempting to meet SACS or WASC institutional effectiveness standards. Indeed, a clearly apparent negative externality of such requirements is a tendency for institutions to back off on important goals that they have no means of assessing. An equally compelling negative consequence is to spread extremely limited available resources for assessment across large numbers of goals — a strategy that may merely result in a lot of very bad evidence.

How can institutions remain true to the spirit of demonstrating effectiveness without falling into the common trap of "measuring everything that moves?" And how can the accreditation process help them to recognize and deal with this problem? One line of attack suggested by emerging experience is to appropriately postpone the discussion of assessment techniques until after a thorough analysis of goals is accomplished. Often, institutions have not thought through the process of systematizing and organizing their goals; many discover that they have large numbers of *de facto* institutional "goals" scattered across different documents and intended for different purposes. Self-study provides an important occasion to sort through these statements in search of the few key "thematic goals" that effectively describe the institution and its aspirations. Often, these are the areas in which the institution sees itself as special or distinctive, and which consequently should be singled out for analysis. An intriguing possibility here is to establish a more active early dialogue between institution and accrediting agency about goals. As a long-term prelude to self-study, for instance, an institution might be requested to work with agency staff and visiting team members to determine well in advance of the team visit a limited and mutually agreed-upon set of goals around which evidence of effectiveness will be organized in appropriate portions of the self-study.

2. ***How to arrive at reasonable expectations regarding outcomes assessment in the process itself.*** At present, the real place and consequences of outcomes assessment in the accreditation process seems murky to many institutional participants. After four years of operation, for instance, institutional reactions to the self-study process under SACS institutional effectiveness standards range from perceptions of "no real change" to reports that the new process in no way resembled the old. Partly this is a result of real variability—particularly in the manner in which individual visiting teams choose to recognize and emphasize assessment. In all likelihood, as experience with assessment develops, variability of this kind will diminish, but the phenomenon does point to some real needs for additional professional development on the part of visiting team members. Current measures to address this difficulty—for example, use of agency staff to cover the "assessment" role, or designation of a specially-chosen team member to address institutional effectiveness—may be effective in the short run. But such practices run the risk of signalling to institutions that assessment is a separate, specialized activity, and may have precisely the opposite of their intended effect in creating a real "culture of evidence."

Similarly, both institutions and agencies must be deterred from slipping into assessment as a compliance exercise—a phenomenon best described by the commonly-voiced institutional complaint: "just tell us what you want us to do and let us do it." Expectations for assessment must be clearly focused on the process itself, not on its particular manifestations or results. Again, this might be clarified at a far earlier in the process than generally now occurs—perhaps in early dialogue among the team, agency staff, and institutional self-study committee. Again, one option here is to ensure early contact between the institution and visiting team to explicitly negotiate the goals and assessment processes which will serve as the focus for demonstrating "institutional effectiveness" in the self-study. At the same time, every effort should be directed toward strengthening the language of current guidelines regarding assessment to better emphasize indicators of good process and the use of results rather than the mere presence of technically sound assessment evidence. As noted, the Malcolm Baldrige National Quality Award criteria currently used to recognize exemplary practice in business and industry provide a suggestive model.

3. ***How to best communicate the results of assessment.*** Once it enters a process like accreditation, the results of assessment are automatically public. Regardless of the care with which it is collected, analyzed and communicated, a particular piece of information can readily be removed from its appropriate context and be used for other purposes. As a consequence, institutions will be understandably wary of sharing bad news, even in a supportive, peer-oriented process. This problem, of course, is not new to accreditation and has been typical of any reported weakness or deficiency. But information about outcomes has the reputation of being far more publicly volatile than traditional information about resources or processes. Partly, of course, institutional fears in this regard are a matter of lack of experience, and there is little evidence from long-established state-based

assessment mandates that institutional reputations suffer major damage from such disclosure. But partly it is a real issue—particularly in the area of student learning outcomes. Nor is the problem solved by concentrating the primary attention of accreditation on the adequacy of the assessment process alone rather than its results. As amply demonstrated by past state-level assessment experience, the only way that an institution can credibly demonstrate the presence of an adequate assessment process is to display a sample of its results.

One potential way to address this difficulty is to avoid putting assessment results all in one place. Reporting evidence of institutional effectiveness under a separate standard, despite its other advantages, can lead to reading this section of the self-study out of context. Similarly, institutions commonly report on such assessment activities as alumni surveys and student satisfaction studies in separate and self-contained parts of a self-study narrative, rather than weaving obtained results into appropriate areas throughout the document. At present, however, institutions receive little guidance about the desirability and benefits of consistently reporting assessment evidence in this fashion. In parallel, experience in state-based assessment processes convincingly suggests that all reported assessment results be accompanied by institutional proposals regarding their implications and what will be done. The admonition "beware the unaccompanied number" applies equally well in accreditation as in state-based accountability reporting. While clearly a key component of accreditation philosophy regarding assessment, and visibly present in agency self-study guidelines, the principle of using obtained results to make appropriate changes has yet to enter the institutional bloodstream in conducting self-study. It should therefore be made a more explicit focus of both team training and of evolving self-study workshops.

4. ***How to get institutions to take the process seriously.*** Again, this is an age-old problem for accreditation—particularly for institutions of large size or established reputation that feel they do not "need" the process. Partly it is a result of uncertain consequences: the role of outcomes in accreditation is sufficiently new that many institutions are unsure of the seriousness with which the agencies themselves will take the process. And indeed, accrediting agencies are faced with a major dilemma in this respect, shared by state agencies, because the sanctions at their disposal tend to be "all or nothing."

Obviously, the only way in which institutions can be induced to take assessment seriously is to convincingly demonstrate that there is something at stake for them in doing so. Several incentives are available, both positive and negative; in one way or another most are now being actively pursued by accrediting agencies. Given the growing salience of state and federal assessment initiatives, one positive incentive is to convince institutions that meeting accreditation requirements will allow them to develop adequate assessment procedures to meet this challenge gradually and in a far more benign environment. Another is to better communicate examples of how institutions have used assessment information to help garner reputation and resources. The major negative sanction practically available to accreditors is simply not to let the institutions alone on the topic if it is not adequately addressed. Interim visits and reports, follow-up questions, and similar mechanisms appear increasingly to be used in this respect and institutions are induced to pay greater attention, if only to avoid this costly state of affairs. At bottom, however, the fact that institutional accreditation as a whole lacks intermediate sanctions is a major problem. Rather than the current "all or nothing" outcome, some form of graduated recognition of institutional merit through accreditation would help to address many of these difficulties, though it would undoubtedly also give rise to a host of others.

5. ***How to sustain institutional action and attention to assessment over time.*** Many have previously lamented the "episodic" quality of institutional accreditation. What makes it a particular problem in the context of outcomes assessment is that to be meaningful, assessment processes require several years of evolutionary development and a significant change in institutional culture. They cannot be initiated or, more importantly, be reactivated at a moment's notice. Those evidence-gathering processes that are developed quickly are immediately apparent to outside reviewers in ways not true for enrollment or financial reporting.

One direct way to address this difficulty is to explicitly require and to look for multi-year assessment results in accreditation. In many cases, the kinds of data-gathering instruments and approaches used to collect outcomes information in connection with self-study are one-time, self-contained events—for example, major surveys or specially-administered general education examinations. While not necessarily discouraging such approaches where appropriate, self-study guidelines might emphasize the provision of more regularly-collected, less extensive efforts. Another point of attack, already mentioned, is to place greater emphasis on providing evidence of how outcomes information has been used to make improvements. This is not a process that can occur overnight, nor can it be improvised quickly to meet the needs of accreditation alone. Real utilization requires data to have been collected well in advance, and an administrative follow-up procedure in place to ensure that resulting improvement issues are kept salient. Required intermediate reports between self-studies and possible targeted follow-up visits or reports to address particular questions are other potential avenues for converting "episodic" assessment into a process more consistent with continuous improvement.

6. ***How to link accreditation's role in outcomes assessment with those of other processes.*** As noted throughout, outcomes assessment is becoming a pervasive process throughout postsecondary education. More than two- thirds of the states, for instance, now require institutions to engage in learning assessment processes substantially similar in kind to those evolving in institutional accreditation. How is assessment through accreditation to be kept distinctive in such an environment and what is its unique contribution? If no answer to this question is forthcoming, institutions will be far more interested in investing scarce energy and resources in complying with government-mandated procedures that have greater consequences.

A particular problem here is how to reconcile the seemingly similar assessment procedures required by different external authorities—often expressed in quite different languages. In substance, for instance, the kinds of evidence required to demonstrate student achievement by the guidelines of the State Council on Higher Education in Virginia (SCHEV) are not markedly different from SACS institutional effectiveness criteria. Yet many institutions have complained that they must engage in "duplicative" data generation and reporting. This suggests first that accrediting agency staff and visiting teams be specifically aware of the kinds of assessment activities and reporting required by applicable state governments. It also suggests that agency staffs maintain much closer liaison with state coordinating or governing board staffs than has typically been the case. Both accreditation and state authorities have an interest in keeping assessment meaningful and avoiding unnecessary duplications of process. At the same time institutional accreditation may provide a unique contribution by emphasizing peer validation of each institution's process of quality improvement. Government authorities are particularly suited for (and are often explicitly charged with) directly establishing the adequacy of obtained results and their match with public purposes. By their very nature, they are less able to examine and stimulate vital ongoing quality-improvement processes. Through its existing strengths of self-study and peer review, institutional accreditation might beneficially concentrate its efforts on this important, but hard to sustain, external role in assessment.

Some Recommended Principles for Future Policy

The many conceptual gaps and unanswered questions occurring throughout this document bear witness to the fact that much learning about the appropriate linkages between assessment and institutional accreditation remains to be accomplished. Up to now, for a variety of justifiable reasons to be sure, the institutional accreditation community has been largely reactive to assessment. States and other government authorities have seized the "high ground" of external accountability with respect to the issue, and some have argued that they allow little left for accreditation to do. Time is indeed short for this condition to be reversed, and certainly all regional agencies are actively responding to the challenge. But if it is to be meaningful or effective, accreditation's response must also be distinctive and useful. This paper provides no remedies. But it does suggest that the following summary principles be used for guidance in developing a consistent policy on assessment in the institutional accreditation process.

- *Assessment should not constitute a "component" of institutional accreditation, but should rather permeate the entire accreditation process.* While guidelines and team training should emphasize the development of techniques for gathering an presenting evidence of institutional effectiveness as a part of self-study, the argument for effectiveness (and its accompanying evidence) should be visible throughout the process. Just as a successful college graduate should be a "lifelong learner", the presence and use of assessment should demonstrate that the institution embodies a "culture of evidence" that enables it to continuously examine and manage itself responsibly, both now and in the future.
- *The primary focus of assessment should be on the effectiveness of the institution as a whole.* Institutional accreditation represents one of the few available opportunities to examine colleges and universities as functioning academic communities. While individual units, departments, and disciplines are all important in demonstrating effectiveness, the focus of assessment in self-study should chiefly be placed upon the attainment of institutional goals. Serious questions should be raised when an institution submits as evidence of its overall effectiveness a case founded only on the individual products of its constituent units or disciplines. Similar questions should be raised about cases that rest only on the fact that the institution currently has in place assessment mechanisms whose sole purpose is to certify individual student achievement.
- *A required focus of assessment should be placed upon student learning.* Institutional effectiveness is a far more inclusive concept for colleges and universities than simply the attainment of claimed learning outcomes. But student learning is the single common goal of the academic enterprise. As a result, the use of assessment in institutional accreditation should visibly center on the presence, appropriateness, and ultimate attainment of specific student learning objectives. Consistent with the first principle, moreover, a concern with student learning should be a visible part of the institution's goals, and should permeate the self-study itself.
- *The primary use of assessment evidence in self-study should be to indicate progress, not to supply definitive judgments about quality.* Assessment, like self-study itself, is often more valuable for the questions that the

process raises than for the concrete results that it may yield. At the same time, no known assessment technique is sufficiently precise to provide definitive answers about institutional goal attainment. For both reasons, the process cannot appropriately be used to determine exactly where a given college or university stands against fixed standards of "quality". Its chief utility instead is to provide credible evidence that the institution is broadly able to attain the goals that it publicly proclaims. Following this logic, guidelines should emphasize the use of multiple indicators and measures in assessment, and that avoid using assessment results judgmentally as the sole basis for decisionmaking.

- *The primary focus of assessment should be the degree to which the process is directed toward and results in continuous improvement.* Assessment is most valuable in the institutional accreditation process when it concentrates on the manner in which obtained results are used. The objective should not be a one-time demonstration of goal-attainment, intended solely to certify a given level of quality. Other organizations—particularly special and professional accreditation and public regulatory bodies—are both better suited to and are more visible in this role. Assessment should rather be used to demonstrate that the institution is willing and equipped to examine itself on an ongoing basis, and is committed to using the results of such examination to actively improve quality on an ongoing basis.

Use of these principles to develop a coherent approach to assessment in the accreditation process might help better define an evolving role for institutional accreditation in a rapidly changing national context. By emphasizing accreditation's role in providing a responsible and visible peer review of the adequacy of each institution's mechanisms for quality assurance, the process may acquire a vital continuing role in the academic community. In increasingly uncertain times, it appears a role well worth exploring further.

*__Peter T. Ewell__ is Senior Associate of the National Center for Higher Education Management Systems (NCHEMS).

Incorporating Assessment into the Practice of Accreditation: A Preliminary Report

Ralph A. Wolff*

Design of the Survey

One of the responsibilities of the COPA Task Force on Institutional Effectiveness is to summarize "the current and prospective use of outcomes assessment and analysis by institutional accrediting bodies in 1) accreditation standards, 2) accreditation guidelines, and 3) such accreditation practices as (a) self study, (b) practices of visiting teams, and (c) deliberations and actions of accrediting bodies. COPA staff have surveyed all institutional accrediting bodies and have compiled their accrediting standards relating to assessment and institutional effectiveness. To further the accomplishment of this charge, the Task Force called for a survey of accrediting agencies to explore the effectiveness in practice of assessment-related standards by institutional accrediting bodies.

A survey was developed which focused on several broad areas: 1) the personal knowledge and experience of the staff member within the accrediting body primarily responsible for assessment-related issues; 2) the knowledge and experience of the accrediting association's staff and accrediting commission with assessment; 3) information on how assessment is defined by the accrediting body; 4) how assessment issues are incorporated into evaluation team training, institutional self studies and commission decision making; 5) personal views of the respondents on barriers and inducements to engage in assessment; and 6) areas of desired additional support.

The questionnaire was sent to the executive director of each of the fifteen institutional accrediting bodies recognized by COPA. The executive director was asked to have the staff member responsible for assessment within the association to complete the questionnaire. Questions were first asked about the personal background and experience of the respondent, to understand whether assessment initiatives with accrediting associations were being guided by staff with considerable experience in the area. Further, information was gathered on the source on the staff member's experience: from readings, attendance at conferences, work with institutions on assessment, or participation in the development of the association's assessment standards on assessment.

Respondents were then asked to evaluate the knowledge and experience of other staff members with assessment. These questions were asked to understand whether there was perceived to be a significant difference between the staff member assigned responsibility for assessment and other association staff, as well as to determine if special training needs existed for association staff.

Questions about the standards and guidelines of each of the accrediting associations were designed to go beyond the collection of assessment standards and criteria already undertaken by COPA staff. Questions were directed to learn how each association defined assessment and to determine if there was any significant variation in this definition. Questions were also asked to determine how recently standards on assessment were adopted or revised, and if the accrediting body followed up the adoption of such standards with special resource materials and workshops.

Sets of questions were asked about the application of the associations standards on assessment. They ranged from an evaluation of the overall importance of the assessment standards in relation to other standards applied by the accrediting body, to questions on how the assessment standards were addressed in team training, institutional self studies and in the actual deliberative processes and decisions of the accrediting commissions.

In a section on personal views, respondents were asked to express opinions about the chief barriers to the greater implementation of assessment within the institutions and the respondent's own accrediting association. Opinions were also sought about the reasons both institutions and their accrediting body have given greater emphasis to assessment. Finally, respondents were asked what additional support they would like to receive in the area of assessment.

Overall, the survey was designed to elicit, at least from the perspective of the staff member responsible for assessment, views about the status of assessment within each institutional accrediting body.

Survey Results

The survey was distributed to all fifteen institutional accrediting bodies recognized by COPA. Responses were received from thirteen. A copy of the questionnaire with tabulated results is attached to this report. Where a quantitative response was requested, a Leikert scale was used with one (1) reflected a low response and five (5) reflecting a high response.

1. *Respondent's Personal Knowledge and Experience With Assessment*

At the outset, it should be indicated that there is no reference point to determine whether the responses to any question was good, in an absolute sense, or not. For example, at what level should we consider the perceived knowledge of staff about assessment as good? A mean score of 3 (out of 5), or 4? Thus while the responses were revealing, there was no external standard against which to judge them.

All thirteen respondents rated their personal knowledge of assessment to be moderate to high with a mean response of 3.6. Only one respondent rated his or her knowledge to be high, seven as moderately high and five as moderate. It is not clear if the five who rated their background as moderate were being modest in light of the increasing complexity and depth of the field of assessment or if they felt that this was not a significant area of expertise.

Questions about the source of the respondents' background on assessment revealed considerable experience by the staff members responsible for this issue. All thirteen respondents indicated that their background with assessment consisted of moderate to extensive readings in the field, with a mean score of 3.8, with eight of the thirteen indicating that they had read widely. Nine of the thirteen respondents indicated they had attended four or more conferences on assessment, which is significant since so much of the assessment movement has grown through the use of conferences. All but one of the respondents participated in the formulation of their association's standards on assessment, and presumably the one respondent who did not was employed after that association's standards were adopted. Ten of the respondents have had moderate to extensive experience with assessment activities at institutions, and several staff members indicated they have worked extensively with institutions in implementing assessment activities (mean score 3.3). All thirteen respondents have participated significantly in the assessment activities conducted by their own accrediting association (mean score 4.3).

2. *Knowledge and Experience of Association Staff and Commission Members*

Only twelve responded to this set of questions, because the thirteenth association did not have additional staff. Ratings by the twelve respondents of their fellow staff members general knowledge of assessment was quite high, though not as high as their own. The mean score for fellow staff was 3.33, above a rating of moderate (3), whereas the mean score of the respondent's self assessment of their knowledge was 3.6. There was a significant difference in the perceived depth of background of the other staff members, however. As to readings, respondents rated their fellow staff at a mean score of 3.0 (moderate), whereas their own readings was rated at 3.8. Similarly, other staff had attended fewer conferences and had participated to a much lesser extent in the formulation of their association's standards on assessment — a mean score of 3.0 compared to 4.6. Interestingly, the rating of other staff members' participation in assessment activities at institutions was rated higher than the respondents' — a mean score of 3.5 compared to 3.3 for the respondents. Positively, there appears to be high participation of other staff in the assessment activities conducted by the association (mean score of 3.92).

The overall knowledge and experience of the respondent's accrediting commission was slightly lower than either their fellow staff based on the mean score (3.30 vs. 3.33). The modal distribution reflects a different picture, however. Six of the respondents rated their commissions to have moderate experience or less, and none rated their commission as high. This distribution reflects a judgment that the level of background and experience of *commissions* with assessment issues is not as well established as commission staffs.

3. *Standards and Guidelines on Assessment*

All but one of the responding accrediting agencies have modified their standards dealing with assessment in the last five years. One indicated that their standards have just been revised and will go into effect in 1992, and two agencies indicated that there have been two revisions in the past five years. Eleven of the twelve indicated that the revisions made the standards more rigorous, while the twelfth characterized the change as more "explicit."

Seventy-seven per cent (77%) of the respondents (ten of thirteen) indicated that special resource materials have been developed to assist institutions address assessment and accreditation expectations. One more indicated that such materials are in development. Similarly, 85% (eleven of thirteen) associations have sponsored special workshops on assessment for their institutions in the last five years.

Defining assessment proved to be more elusive. Of seven different areas encompassed by "assessment" all thirteen respondents agreed that their association's definition of assessment included "accomplishment of *institutional* goals and purposes." Twelve respondents agreed that assessment also included two other areas: evidence of program success and linkage to institutional planning efforts. Eleven respondents indicated that assessment included evidence of student learning, which seems consistent with the thrust of the national assessment movement, and ten indicated that assessment was to be linked to program review processes. The least agreement was found on assessment including "evidence of the effectiveness of faculty and student research" (four respondents), "evidence of building a multicultural community" (five respondents), and "evidence of the effectiveness of the continuing education program" (six respondents).

Importantly, the importance given to the association's standards on assessment in relation to all standards was quite high, being given a mean score of 3.85, with twelve respondents rating the importance to be moderate (3) or above.

Team Training Workshops

Seven associations (64%) of the eleven responding indicated that they always include assessment issues in their team training workshops, while one reported never doing so (mean score 4.18). In such workshops, respondents were asked the proportion of time devoted to assessment. Only six responded to this question, ranging from a high of 50% to a low of 10%, averaging at 29%. In these workshops, all those responding indicate that assessment is discussed in the context of all accrediting standards.

Self Studies

Questions about how institutions treat assessment in their self studies yielded interesting results. The respondents were asked to distinguish between the treatment of assessment which *described* activities, with those that *analyzed data* in the self study based on assessment, and those that identified *improvements made from assessment activity*. On average, the mean scores of these three questions did not differ significantly. Eleven of twelve respondents indicated that most institutions described assessment activities (mean score 3.58). A similar distribution of responses occurred about self studies analyzing data (mean score 3.5). Not surprisingly, the mean score for institutions identifying improvements in their self studies was significantly lower, at 2.91, and five respondents indicated that institutions rarely identified improvements from assessment activities in their self studies.

Respondents were asked to venture a personal opinion as to the percentage of institutions accredited by their association in which assessment was a serious enterprise. The responses reflected a surprising and widely disparate range. One respondent indicated all, another 90%, but explained that the institutions accredited were mostly public institutions. The lowest estimate was 25%. This seems at odds with perceptions of those familiar with institutional responses to state mandated assessment that institutional responsiveness is quite limited. Perhaps the question was perceived differently by those responding.

Team Visits

Several questions were asked about how assessment is treated in the composition of the evaluation team and in the evaluation team report. Accrediting teams do not always have at least one team member with background and experience in assessment for comprehensive visits. Only one association always has a team member with an assessment background, and one association never does and another rarely has such a team member. Ten associations sometimes do (mean rating 3.23). This suggests that the application of the accrediting association's standards on assessment in the course of a site visit might be limited by the lack of at least one person on the team with considerable experience in the areas. When asked if there is a discussion of assessment in team reports, the same agencies that never or rarely have a team member with an assessment background report that assessment is rarely or never discussed in the reports. The other agencies reported that team reports include discussions of assessment at a moderate level, slightly better than the rate at which team members with assessment backgrounds are included on teams (mean score 3.46). It might be assumed that even without an assessment team member, the team is aware of commission standards and includes some information about assessment in the reports. Still, it would appear that inclusion of assessment in team reports is still not consistently practiced.

For visits other than comprehensive, assessment is included less frequently, which would be expected. Special visits typically focus on specific concerns identified by the accrediting commission or a previous team, and assessment would not be an focus of the visit unless it was one of the specific concerns. Four (4) associations reported that assessment is rarely or infrequently included in special team reports, and the mean response was 2.92.

On the other hand, assessment is the basis for major team recommendations, with nearly all associations reporting that major recommendations are sometimes included in team reports. Major recommendations are given serious consideration by institutions since institutional responses to such recommendations are typically reviewed by the subsequent visiting team. (Mean response 3.23)

Commission Actions

Respondents were asked if their accrediting commission discusses assessment when reviewing team reports. Only one association indicated that assessment is always considered when action is taken following a comprehensive evaluation. Overwhelmingly, respondents indicated that assessment is discussed when it is raised by the evaluation team, but much less frequently is assessment raised by the accrediting commission on its own. A similar pattern emerges regarding whether accrediting commissions discuss assessment when taking action following other types of visits. It is important to note that special reports or visits involving assessment issues are now being required by most commissions, most frequently when raised by the evaluation team, but also occasionally by the commission on its own.

Interestingly, when asked what weight is given to assessment in making accrediting decisions, respondents reported a wide range of commission practices. Two respondents indicated the weight given to assessment is low, whereas ten respondents indicated that the weight given is moderate or better (mean response 3.25)

4. *Personal Views of Respondents*

Respondents were asked to identify three chief barriers to the implementation of assessment within institutions. The most commonly cited barrier (ten respondents) is the lack of knowledge within institutions of how to conduct assessment as well as an understanding within the institution that assessment can lead to improvement. Other frequently identified barriers included: lack of commitment from the chief administrative leaders of the institution (four respondents), lack of financial resources (four respondents), faculty resistance (three respondents), lack of adequate instruments or techniques of evaluation (three respondents).

Respondents were also asked to list three barriers to greater implementation of assessment *within their own accrediting association*. A wide range of responses were provided. The most frequently cited item (six respondents) was a lack of understanding of assessment, a similar response to the previous question dealing with institutional implementation of assessment. Other responses included: lack of commitment from association leadership, lack of resources committed to provide workshops, fear of negative results or what assessment might reveal, staff unwillingness to change, a lack of definition of what is acceptable assessment, and a concern about requiring institutions to invest heavily in assessment when financial resources are already so strained. One respondent indicated there are no barriers within that person's association, but it should be noted that this respondent also indicated that the association had not yet adopted formal standards on assessment nor developed resource materials on the topic.

Respondents were next asked what has led institutions served by their association to undertake assessment. Interestingly, the leading factors cited are largely pressures external to the institution, whether from the accrediting association itself (ten respondents), states (five respondents), or a broader notion of "public accountability" including pressure from the institution's constituencies and sponsoring denomination(s) (seven respondents). Seven respondents also indicated that institutions do engage in assessment to learn about themselves or to improve, but these comments were cited along with citations of external pressures.

Finally, respondents were asked their opinions about what led their own agency to incorporate assessment as part of the accrediting process. Three respondents emphasized that assessment is not new to accreditation, notwithstanding all of the larger public concerns. The most frequently cited factor was the new assessment requirement of the Department of Education (five respondents), followed by more general statements of public accountability (four respondents), COPA recognition requirements (three respondents), and a philosophical desire to focus less on inputs and resources (two respondents).

Additional Support Desired

Respondents were asked what additional types of support they would like to receive in the area of assessment. Eleven respondents indicated they would like to have workshops address specific assessment techniques; eight asked for a workshop on how assessment is used by other accrediting agencies; and seven requested suggested readings. In the open comment section, several respondents asked specifically for dialogues on important issues: how to train evaluators in the area of assessment, how to determine the acceptability of evaluation team findings in the area of assessment, and how to build better understanding and consent on the part of accrediting commissions so that assessment can be more effectively incorporated into the decision making process.

Analysis and Conclusions

What is clear from the survey responses, is that the respondents generally take great pride in the ways in which their agencies incorporate assessment into the accreditation process. From the responses provided, several major findings appear warranted:

1. The accreditation staff responsible for assessment have considerable experience with assessment, by way of readings, conference attendance, some participation in assessment activities at institutions, and extensive participation in the formulation of their agencies' own standards on assessment. While it is not clear whether these staff would be the ones attending a COPA workshop on assessment, if they are, a workshop could run the risk of "preaching to the converted." Perhaps special attention will need to be paid as to the constituency desired to attend COPA workshops, and then tailor the material to the level of experience of those attending.

2. The background and experience of other association staff is reasonably good, although other staff have less familiarity with assessment through readings and conference attendance. One issue that emerges is whether the COPA assessment project should not target training materials for other staff than the person primarily responsible for assessment within the association. Indeed, it might be advisable to consider "training" the primary assessment staff person to lead workshops within their own associations for their fellow staff members and commission members.

3. Respondents rated the overall knowledge and experience with assessment of their accrediting commission to be moderate. Yet a lack of understanding of assessment by the commission was the most frequently cited barrier to further implementation of assessment within the accrediting body. Special attention should be paid to commissioner experience and understanding of assessment. Respondents also reported that discussions of assessment at the time accrediting decisions are made by the accrediting commission are most frequently raised only when an issue relating to assessment arises in the evaluation team report. This would seem to suggest that a fruitful area of further attention is how to improve the means by which assessment is incorporated in the actual decision making process.

4. Twelve of the thirteen accrediting associations responding to the survey indicated that they had modified their standards relating to assessment in the past five years, and these modifications have made the association's expectations more rigorous. This is an important statement about the responsiveness of the accrediting community to the issue and should be given wider visibility. Further, there has been a

widespread effort on the part of accrediting associations to develop resource materials and special workshops for institutions on assessment issues. This is a new role for accrediting associations — both in terms of adopting standards on a leading issue somewhat ahead of institutional practice, and then supporting institutions to move more fully in this area. As a result, accrediting associations are moving away from an exclusively evaluative role, but are consciously working to lead institutions toward better assessment practices. This shift in role is worthy of further discussion and comment.

5. There is some variation of the ways in which the accrediting associations surveyed define assessment. This is not surprising given how recently most associations have adopted standards on assessment and how many definitions there are for assessment nationally. Nonetheless, there is a need for institutions and team members to understand how their accrediting association defines assessment and expects such a definition applied. Further discussion would seem warranted within accrediting associations and between associations on how assessment is defined, and what must be included or may be.

6. Responses to questions on the quality of self study reports dealing assessment indicate that institutions are better able to describe assessment activity than identify improvements which have come about as a result of assessment. Further attention should be paid to a) how different accrediting associations expect institutions to report and analyze assessment activity, and b) giving wider dissemination to examples of improvements brought about by assessment to assist in institutional understanding of how assessment can benefit the institution.

7. The requests for additional support from the respondents seem to indicate that the greatest need is a better understanding of specific assessment techniques. Based on open-ended comments, this would seem to mean there is an interest to know how specific techniques or methodologies can be used to improve quality at the particular range of institutions served by the accrediting associations surveyed.

8. Finally, there seems to be a need to address in workshops how teams should evaluate institutional assessment efforts as well as how commissions should review team findings about assessment.

A postscript: One respondent, in the open comment section to the survey, observed that "outcomes assessment is an over-reaction to process oriented accreditation. The pendulum has swung too far — assessment is not a solution to measuring educational quality. . . . Hopefully, assessment will just fade away and become another tool in the accreditor's toolkit." All of the respondents indicated that when assessment is introduced in team training workshops, there are discussions about how to view assessment in the context of the overall evaluation process. No matter what we do, collectively and individually in our accrediting associations, it will always be important to remember that assessment is a *means* and not an end in and of itself.

***Ralph A. Wolff** is the Associate Executive Director of the Accrediting Commission for Senior Colleges and Universities of the Western Association of Schools and Colleges.

TABULATED RESPONSES

Survey of Accreditation and Assessment
COPA Task Force on Institutional Effectiveness ◆ July 1991

MR = Mean Response

A. Personal Knowledge and Experience with Assessment

1. Please rate your own knowledge of assessment.

Responses		*5*	*7*	*1*		*MR = 3.6*
	1	2	3	4	5	
	Low		**Moderate**		**High**	

2. My background and experience with assessment consist of the following:

a. Readings

Responses			*5*	*5*	*3*	*MR = 3.8*
	1	2	3	4	5	
	None		**Moderate**		**Extensive**	

b. Attendance at conferences
[Circle number attended]

Responses		*1*	*1*	*1*	*5*	*1*	*3*
	0	1	2	3	4	5	5+

c. Participation in the formulation of my association's standards on assessment

Responses		*1*		*2*	*10*	*MR = 4.6*
	1	2	3	4	5	
	None		**Moderate**		**Extensive**	

d. Participation in assessment activities at institutions

Responses		*3*	*5*	*3*	*2*	*MR = 3.3*
	1	2	3	4	5	
	None		**Moderate**		**Extensive**	

e. Participation in assessment activities conducted by my association

Responses			*3*	*3*	*7*	*MR = 4.3*
	1	2	3	4	5	
	None		**Moderate**		**Extensive**	

f. Other (please explain) ______________________________

____

B. Knowledge and Experience of Association Staff and Commission Members

1. I would rate the general knowledge of other association staff about assessment as

Responses		*1*	*7*	*3*	*1*	*MR = 3.33*
	1	2	3	4	5	
	Low		**Moderate**		**High**	

2. The background and experience of other staff members consist of the following:

a. Readings

Responses		*4*	*4*	*4*		*MR = 3.00*
	1	2	3	4	5	
	None		**Moderate**		**Extensive**	

b. Attendance at conferences

Responses	*1*	*5*	*3*	*3*		*MR = 2.66*
	1	2	3	4	5	
	None		**Some**		**Many**	

c. Participation in the formulation of my association's standards on assessment

Responses	*1*		*6*	*3*	*1*	*MR = 3.00*
	1	2	3	4	5	
	None		**Moderate**		**Extensive**	

d. Participation in assessment activities at institutions

Responses		*1*	*6*	*3*	*2*	*MR = 3.5*
	1	2	3	4	5	
	None		**Moderate**		**Extensive**	

e. Participation in assessment activities conducted by my association

Responses		*1*	*3*	*4*	*4*	*MR = 3.92*
	1	2	3	4	5	
	None		**Moderate**		**Extensive**	

f. Other (please explain) ______________________________

3. I would rate the overall knowledge and experience of my accrediting commission with assessment to be

Responses		*3*	*3*	*7*		*MR = 3.30*
	1	2	3	4	5	
	Low		**Moderate**		**High**	

C. Standards and Guidelines on Assessment

1. My accrediting association has modified its accrediting standards or criteria in the last five years to address assessment.

1987 (3), 1988, 1989 (3), 1990, 1992
Responses *12* Yes *1* No If yes, specify date: *1984 & 1991, and 1989 & 1991.*

2. If the answer to #1 above was yes, what was the net effect of these changes on those accrediting standards dealing with assessment?

_____ Became less rigorous
_____ Stayed about the same
Responses *11* Became more rigorous
Responses *1* *Became more explicit*

Comments: ____________________

3. In the last five years, my accrediting association has developed special resource materials to help institutions address assessment and accreditation expectations.

1987 & 1989 (2), 1989, 1990 (4)
Responses *10* Yes *3* No If yes, specify date(s): *1990-91*

4. My accrediting association has developed special workshops on assessment for *institutions* in the last five years.

Annually since 1988 (3),
Responses *11* Yes *2* No If yes, specify date(s): *1990 (3), 1991, April-May*
1991 & annual meeting

5. Assessment in my accrediting association is specifically defined to include: [Check all that apply.]

13 Accomplishment of *institutional* goals and purposes
11 Evidence of student learning
8 Evidence of teaching effectiveness
12 Evidence of program success
5 Evidence of building a multicultural community
4 Evidence of the effectiveness of faculty and student research
6 Evidence of the effectiveness of continuing education
10 Linkage to program review processes
12 Linkage to institutional planning efforts

6. The overall importance of my agency's standards on assessment in relation to all accrediting standards is

Responses		*1*	*4*	*4*	*4*	*MR = 3.85*
	1	2	3	4	5	
	Low		**Moderate**		**High**	

D. Team Training

1. Assessment is included as a part of team training workshops.

Responses	*1*		*2*	*1*	*7*	*MR = 4.18*
	1	2	3	4	5	
	Not at all	**Sometimes**			**Always**	

2. The proportion of time devoted to assessment in team training workshops is roughly _____*%. 10%, 15%, 20% (2), 30%, 50%, and increased since 1990*

3. The focus of training about assessment is directed toward:
[Check all that apply.]

8 Standards on assessment
8 Guidelines to interpret assessment standards
8 Different types of assessment methods institutions might undertake
11 How assessment should be viewed in the context of the total evaluation process

E. Applications of Assessment in the Accrediting Process

SELF STUDIES

1. In operational terms, self studies currently being submitted *describe* assessment activities.

Responses		*1*	*3*	*4*	*4*	*MR = 3.58*
	1	2	3	4	5	
	Not at all			**Extensively**		

2. Self studies currently being submitted display *analysis* of data gathered by assessment activities.

Responses		*1*	*6*	*3*	*2*	*MR = 3.5*
	1	2	3	4	5	
	Not at all			**Extensively**		

3. Self studies currently being submitted identify *improvements* made in the institution as a result of assessment.

Responses		*5*	*4*	*2*	*1*	*MR = 2.91*
	1	2	3	4	5	
	Not at all			**Extensively**		

4. In my personal opinion, assessment is a serious enterprise in _____% of the institutions my agency accredits. *25%, 30% (2), 33%, 40%, 50%, 70% and growing, 80%, 90% (public institution), all*

TEAM VISITS

1. At least one team member with background and experience in assessment is included on comprehensive evaluation teams.

Responses	*1*	*1*	*6*	*4*	*1*	*MR = 3.23*
	1	2	3	4	5	
	Rarely		**Sometimes**		**Always**	

2. A discussion of institutional assessment efforts is typically a part of comprehensive visiting team reports.

Responses	*1*	*1*	*2*	*9*		*MR = 3.46*
	1	2	3	4	5	
	Rarely		**Sometimes**		**Always**	

3. Assessment is included in team reports of other types of visits.

Responses	*2*	*2*	*5*	*3*	*1*	*MR = 2.92*
	1	2	3	4	5	
	Rarely		**Sometimes**		**Always**	

4. Major recommendations about assessment are included in team reports.

Responses		*1*	*8*	*4*		*MR = 3.23*
	1	2	3	4	5	
	Rarely		**Sometimes**		**Always**	

COMMISSION ACTIONS

1. My accrediting commission discusses assessment when it takes action following a comprehensive evaluation.

_____ Never
_*7*_ Sometimes on its own
_*10*_ When it is raised by the evaluation team
_*1*_ Always

2. My accrediting commission discusses assessment on other types of reports.

_____ Never
_*8*_ Sometimes on its own
_*9*_ When it is raised by the evaluation team
_____ Always

3. My accrediting commission requires special reports or visits addressing assessment.

_*1*_ Never
_*4*_ Sometimes on its own
_*8*_ When it is raised by the evaluation team
_*4*_ Regularly

4. The weight given to assessment in the overall accrediting decision is

Responses	*1*	*1*	*5*	*4*	*1*	*MR = 3.25*
	1	2	3	4	5	
	Low		**Moderate**		**High**	

F. Personal Views

1. In my opinion, the three chief barriers to greater implementation of assessment *within institutions* are

Ignorance, lack of presidential and CAO commitment, faculty resistance

Lack of time and staff, lack of knowledge, lack of "from the top" support

An institution not understanding what it wants to do, and knowing how to do it

Legitimate concern about how the outcomes data collected might be comparatively analyzed and (mis-)used; distaste for standardization in a profession that prizes the autonomous teacher and institution; continued confidence in the belief that good inputs assure good outcomes.

Lack of a clearly focused appropriate mission statement, lack of explicitly stated measurable goals and objectives, and lack of local ownership of assessment as a critical element in demonstrating institutional effectiveness

Lack of in house academic expertise in many theological schools, philosophical differences about appropriate outcomes of theological education, and typical small size of seminaries

Faculty resistance, finances, apathy

Lack of Support from the President or Chancellor, faculty resistance to the concept, and lack of understanding/knowledge of the benefits to the institution.

Availability of trained evaluators, mix of subjectivity/objectivity inherent in evaluation process, methodologies that will capture adequate outcomes.

Lack of knowledge regarding assessment technicians, lack of assessment instruments in specialty fields, lack of resources.

Institutional culture, which views assessment with alarm; some institutional leaders are not yet sufficiently committed to provide necessary leadership; and lack of sophistication about evaluation.

Time, knowledge of process and understanding of benefits, and ability to apply results in planning/implementing.

The absence of financial resources to implement assessment programs.

2. The three chief barriers to greater implementation of assessment *within my accrediting association* are

Ignorance, fear of having to require institutions to spend more money, concern over having institutions do something they don't know how to do - being too prescriptive.

Meaningless of assessment in relation to accreditation, lack of understanding, it was forced on us by the government.

There are no barriers.

Staff time that can be devoted to institutional assessment is, perforce, limited; institutional assessment programs are idiosyncratic and, therefore, relatively hard to understand; Commissioners have only moderate amounts of assessment experience: there is no real expert on our Board.

Institutions that traditionally focus on "describing" vs analyzing and implementing change based on data analysis. Lack of evaluation programs and limitation of feedback mechanisms. The Association's/Commission's lack of experience and expertise with assessment activities.

For the most part, I would be inclined to say they reflect the association counterpart of the three issues noted for schools above

We are presently implementing an assessment requirement that calls for all institutions to document student academic achievement.

The nebulous nature of assessment and making qualitative judgments about the planning and evaluation process, lack of definition of what is acceptable, given the variety of institutions which constitute the membership, and unfamiliar territory — not easily evaluated.

Lack of adequate team training (compounded by turnover of team personnel), lack of more specific criteria, and lack of needed assessment instruments.

The apparent belief that 'someone' will develop the definitive system, which can then be adopted with a minimum of pain; competing economic imperatives which prevent some institutions from staffing appropriately; the flurry of demands for *accountability from too many sources with too many agendas in addition to institutional improvement.*

Desire of members not to face realities of shortcomings, fear of consequences of negative actions, lack of deeper understanding/appreciation of assessment benefits.

Staff unwillingness to change methods and reduce to writing past practices that would give evidence of assessment, the unwillingness to use information provided for program enhancement, and the resources necessary to conduct appropriate workshops.

3. The three chief factors leading institutions served by my association to undertake assessment are

The assessment initiative of our Commission, the nationalal assessment movement, and the desire to learn more about themselves

Required by law, required by accrediting body, possible utility in perfecting institutional marketing effort

Commitment to institutional quality and improvement. State mandates (some institutions have not developed that commitment as of yet.)

Pressure from their respective states; pressure from consumers; felt need to address their own failures

Increased emphasis by regional accrediting agency to provide evidence of institutional effectiveness and de-emphasizing self studies which are merely descriptive in nature. Increased emphasis by regional accrediting agency to focus on "useful" self-studies, tied to strategic planning. Public demands for accountability.

Pressure from ATS or one of the regional associations, internal desire of the institution, and pressure from sponsoring denomination

State/government pressures, accreditation requirements, increased accountability

Requirement for gaining and maintaining membership, state requirements, funding.

Types of education lend themselves to assessment/outcomes, need to capture "essence of schools," ability to use evaluation results in working with schools in improving their effectiveness

Desire for self improvement, it is required by self-study process, and demands for accountability on the part of the constituency.

General and growing recognition that the burden of demonstrating effectiveness rests with the institution — defensive reasons exist, in addition to planning imperatives; the recognition that effective planning rests on systematic evaluation; steady pressure from a variety of sources (1) accreditation (we were first), state agencies, and categorical funding.

A sincere desire to improve, understanding benefits of assessment in planning changes, fear of punitive action such as being dropped.

Compliance with standards of membership and willingness to use interpretive guidelines and take steps necessary to be in compliance with standards of membership.

4. The three chief factors leading to assessment being used as part of the accrediting process are

Leadership of SACS, national and federal concerns, and staff leadership

Federal Recognition Criteria, COPA

Part of the definition of an accredited institution is its demonstrated capacity to accomplish its purposes and that it continues to do so. The institution's capacity to accomplish purposes is an indicator of institutional quality. The ability of the institution to plan and evaluate are valid indicators of institutional quality. Assessment has always been a part of accreditation.

Outcomes assessment provides evidence of congruence between stated goals and eventual outcomes; Like all of American higher education, accrediting bodies are moving toward standardized expectation; there is a growing consensus on relatively reliable tests and measures

Clarification and strengthening references to assessment in Accreditation Handbook. Expectation by Commission for institutions to provide evidence of institutional effectiveness as a larger matter than simply learning outcomes. Requests through the Accreditation Handbook, 1988 edition, for "Analysis and Appraisal" and "Evidence of Effectiveness."

Need to guide schools toward better internal decision-making processes, and general concerns in accrediting community.

We have long had assessment; it is not new! There is new emphasis because of: federal requirements, COPA, and greater need for accountability.

Less emphasis on resources and more focus on the use of resources to further the goals of the institution, more emphasis on a clearly defined purpose and articulation of goals, enable review of the non-traditional programs.

Types of education lend themselves to assessment/outcomes, need to capture "essence of schools," ability to use evaluation results in working with schools in improving their effectiveness.

Public demands for accountability, requirements of DE/COPA, and increased use of non traditional delivery systems.

Institutional improvement rests on a knowledge base, effective planning requires assessment, accreditation decisions require more than fuzzy opinions about institutions.

Desire to create standards which result in improvement, new information which shows how to apply assessment techniques in school reviews and evaluations, new criteria and demands made by public agencies and USDE regulations for accreditors.

Institutions have always been involved in assessment at the technical and career level and it is now a matter of reducing past activities to writing, and the recognition that the assessment of institutional effectiveness can lead to future program enhancement.

5. What types of additional support would you like to receive in the area of assessment?

_____ Suggested readings
_____ Workshop on how assessment is used by different accrediting agencies
_____ Workshop on specific assessment techniques
_____ Other (please list below)

What has worked well - models of g.e. review, for example. A survey of good practices.

Outcomes assessment is an over-reaction to process oriented accreditation. The pendulum has swung to far — Assessment is not a solution to measuring educational quality. In most cases, it's a phony exercise in dubious statistics. Hopefully, assessment will fade away, and become just another tool in the accreditor's toolkit.

Opportunities for dialogue among those of us within *the accreditation community.*

Guide or framework for training and evaluating (for example: what evaluative criteria should be brought to bear? how? when?) outcomes assessment evaluators.

Please note, too, our Commission's extensive publications on assessment: Assessment Workbook, Spring 1991, Fall 1990 NCA Quarterly, "Sharpening the Focus on Assessment," Fall 1991 NCA Quarterly will also be on assessment.

Workshop on how to review the acceptability of findings which result from evaluation.

Help in development of instrumentation appropriate to specialized institutions/programs.

The community of regional accrediting agencies should establish a common baseline of minimum expectations for institutional assessment. Literature available is diffuse, sometimes contradictory, and frequently too technology dependent. A well-crafted testimonial policy statement should be succinct, describing systems that are achievable with modest resources, comprehensible by non-researchers, and focused on matters of greatest importance to the teaching-learning enterprise.

Thank you for completing this survey. Please make any additional comments on the reverse side or attach a separate sheet.

Methods for Outcomes Assessment Related to Institutional Accreditation

Thomas P. Hogan*

Preface

This paper was prepared for the Council on Postsecondary Accreditation (COPA) "Outcomes Analysis Project," funded by the Fund for the Improvement of Postsecondary Education (FIPSE). More specifically, the paper was developed for the project's Task Force on Institutional Effectiveness (one of three task forces operative within the project). Among the several topics addressed by this task force, the paper attempts to respond to the call for a guide to the methods and instruments related to outcomes assessment, with an analysis of their strengths and weaknesses—all within the context of institutional effectiveness.

The paper begins by identifying a number of conceptual issues which may have significant influence on the selection of particular methods of assessment, then goes on to an overview of various methods or instruments, with commentary on their strengths and weaknesses. I have aimed the paper principally at those individuals who have considerable experience in higher education but not much experience in assessment and, hence, are feeling overwhelmed or confused by the rapidly developing demands for outcomes assessment; and at those individuals who may have considerable experience in assessment methodology but not in the context of higher education's concern with accreditation matters.

This paper is complemented in important ways by three other papers prepared by members of the Task Force on Institutional Effectiveness. Peter Ewell's "Outcomes Assessment, Institutional Effectiveness, and Accreditation: A Conceptual Exploration" provides an insightful analysis of the role of assessment in the accreditation process. Trudy Banta's "Selected References on Outcomes Assessment in Higher Education: An Annotated Bibliography" provides an excellent overview of the literature on the topics addressed by the Task Force and, in a way, serves as the references for all of the other papers. Ralph Wolff's "Incorporating Assessment into the Practice of Accreditation" provides results of a survey of members of the accrediting community on topics related to the concerns of the Task Force.

While I accept full responsibility for the contents of this paper, I would like to thank the other members of the Task Force on Institutional Effectiveness (Peter Ewell, Trudy Banta, Ralph Wolff, Dorothy Fenwick, James Rogers, the Task Force Chair, and Richard Millard, project coordinator) for helpful comments on and a fruitful discussion of an earlier draft of the paper.

Introduction

A real estate agent confidently tells a customer: "This house is located in the best school district in the area." A high school principal, beaming, reports to a PTA meeting: Our graduates this year have gone on to some of the best universities in the country." A proud uncle tells his golfing partner: "My nephew was just accepted at St. Basil's College; I hear it's the best small college in the tri-state area."

Comments such as these are commonplace. But what can they possibly mean? What basis do people have for formulating these judgments? Or for accepting the judgments rendered by others on these apparently weighty matters? Perhaps more to the point: Can we identify any rational bases for making statements about the quality of educational institutions? If so, can we apply these bases in the formal method we have developed for determining the success of educational institutions: the accreditation process?

Historically, our judgments about the quality of educational institutions have been based principally on what are called *inputs*. A "good" institution was one that paid high salaries to its teachers, had a good physical plant, invested heavily in its library, provided up-to-date laboratories and computer facilities, and so on. An institution which had lower salaries, crowded classrooms, last year's computer models, etc. was thought to be second-rate.

Increasingly in recent years, people have become dissatisfied with, or at least suspicious of, this "input" approach to defining quality. The suspicions arose first, not at the post-secondary levels of education, but at the elementary and secondary levels with the advent of the "accountability movement." This movement had its origins in the latter half of the 1960's, which witnessed massive infusions of federal dollars into educational programs. Tied to these infusions were requirements that the programs be evaluated. In all too many cases, it was noted that the new funding made little or no difference in student learning. In fact, at a macro level, it was noted that, as spending increased, student achievement seemed to deteriorate. Total expenditures—local, state, and federal—continued to increase at a steady clip. But such indices as SAT scores, college-going rates, retention rates, literacy rates, etc. declined. In this milieu, attention turned to the *outputs* of education. The 1970's witnessed an outpouring of "school effectiveness" studies. These studies started by identifying schools which seemed to produce good outcomes (usually defined in terms of student learning as measured by tests), then attempted to determine the "input" or "process" characteristics of the schools. By the 1980's and continuing to the present, it has been an accepted fact at the elementary and secondary school levels that the quality of a school must be defined by its outcomes.

Postsecondary education has come late to this game, but it is tracing the same path as that traversed by lower levels of education. For example, the staple of the accreditation process at the postsecondary level has been an input model. Library resources, classroom space, faculty salaries, lab facilities, etc. have been the key subjects of study. The process has operated on the *assumption* that if the inputs are in place the outcomes will inevitably follow. That assumption has increasingly been called into question, just as it was in an earlier day at the pre-collegiate level. Hence, the postsecondary accreditation process, while clearly not abandoning steady attention to inputs, calls increasingly for *outcomes assessment*. (Specific references to outcomes assessment in accreditation materials is documented in another paper prepared for this COPA project; see *Outcomes Assessment and Analysis: A Reference Document for Accrediting Bodies*, June, 1991; available from the Council on Postsecondary Accreditation for $6.00.) Concern for outcomes assessment has clearly been one of the two or three major topics of discussion in post-secondary education in the past decade. However, the methodology appropriate for outcomes assessment is still somewhat unfamiliar territory for many people in higher education, including those involved in the accreditation process.

The principal objective of this paper is to identify various *methods* which might be useful for outcomes assessment, and to comment on the relative strengths and weaknesses of these methods, particularly in relation to their use for assessing *institutional effectiveness*. However, before moving directly to a treatment of these methods, we must treat a number of other issues. Our experience in working with various assessment methods—and, more especially, in working with various faculty, staff, and administrators in both educational institutions and accrediting bodies—tells us that the strengths and weaknesses of the methods can be discussed fruitfully only after some more general questions are confronted. Further, while many people think that the *real* problems of outcomes assessment are technical or methodological questions, in fact, the really difficult problems lie in the realm of these general questions. So, we proceed first to treat these general questions, then refer to them in the discussion of particular methods.

The First Key Question: Final Status or Value Added. The first key question to be answered, or choice to be made, before deciding about methods to be employed for assessing institutional effectiveness, is whether to concentrate on the *final status* of institutional results or the *value added* by the institution. The names for these two approaches are largely self-explanatory, but let us comment briefly on each.

In the *final status* approach, we are concerned only with the final outcomes evident for an institution's "products." (The "products" may be graduates, community services, faculty research, or whatever else is called for in the institution's catalog of intended outcomes.) It makes no difference what the starting point was. It makes no difference whether the institution contributed to the development of the final status. Perhaps more appropriately, it is simply *assumed* that the institution contributed, at least to a significant extent, to the final status.

In the *value added* approach, we are concerned primarily with the extent to which the outcomes have been influenced by the institution. Are graduates getting better jobs than they would have acquired without attendance at the institution? Has writing ability developed beyond what would have been expected from simple age-related maturation? Do a large percentage of an institution's graduates enter prestigious graduate schools because the institution trains its students exceptionally well, or just because it starts with a very bright group of freshmen?

We should make three observations about the choice between the final status and the value added approaches to outcomes assessment. First, nearly everyone agrees that the value added approach is the one that should be used. Second, despite this agreement, nearly all of the outcomes measures actually used by institutions are of the final status type.

Third, and most important for purposes of this paper, there are clear methodological implications involved in the approach that is adopted. Specifically, the methods of collecting, analyzing, and interpreting information are enormously more complex in the value added approach than in the final status approach. In the final status approach, we deal with just one measurement at a time, e.g. a test score, job placement rate, or index of satisfaction. In the value added approach, we need at least two measurements. The two measurements can occur in either one of two contexts. First, they can come from a "before-and-after" design, i.e. before and after association with the institution. Second, they may come in a "comparison" design, i.e. after association with the institution vs. after non-association with the institution (which in practice may mean association with some other institution or association with no institution at all). Ideally, both designs should be used in a coordinated manner, thus involving a total of *four* measurements.

It is clear that the value added approach, admittedly the preferable one, presents significant challenges[1]. In many ways, one is thrust back on the final status approach, with its attendant inferential uncertainties. Those interested in outcomes assessment related to institutional effectiveness need to be aware of the contrasts presented by these two approaches: both the questions they answer and the methods required to yield their answers.

The Second Key Question: Final Answers or Meaningful Processes. The second key question to be answered or choice to be made relates to what has emerged as an important philosophical—or perhaps more accurately, a stylistic—approach. On the one hand, some maintain that it is possible to develop relatively definitive answers to the questions of whether desired outcomes are being achieved. And, the outcomes assessment should provide these final answers.

On the other hand, some maintain that it is too difficult to provide the final answers, but that does not mean that we should abandon the pursuit altogether. Rather, what is important, according to this position, is that we assure that meaningful processes are established to continually generate and examine information related to outcomes assessment.

Is this a distinction without a difference? That is, do the final-answers and meaningful-processes positions amount to the same thing in terms of accreditation concerns? By no means. There are at least two major differences in the practical consequences of these positions in terms of accreditation reviews.

First, there is a difference in *who* is involved and in *how* they are involved in outcomes assessment activities. In the final-answers approach, it makes little difference who is involved or how they are involved (except to insure the technical correctness of procedures, as per the second issue, considered below). The sole concern is with providing correct, valid answers. According to this position, it is probably best to have one or a few psychometric experts simply provide a package of answers to an institution about its outcomes. Then, an accrediting group can examine the package of answers.

In contrast, according to the meaningful-processes approach, we need to worry about broad involvement of the constituencies affected by the outcomes assessment, both with respect to selection of methods for assessment and interpretation of results. An accreditation review would be no more interested, perhaps even less interested, in the results of the assessment than in how the results were obtained and used. The review would treat, for example, what groups of faculty, staff, and other groups were involved in developing the assessment methodology; how widely the results of the assessment were disseminated and discussed; and similar issues.

The second major difference between the final-answers and meaningful-processes approaches to outcomes assessment relates to the *technical characteristics* of the methods employed. The technical characteristics to which we refer include such matters as reliability and validity of measures, sampling techniques (if sampling is employed), adequacy of normative data (if normative comparisons are involved), and other such matters.

In the final-answers approach, one needs to worry a great deal about such matters. It is critical that the institution bring expertise to bear on these matters. And, it is critical that members of an accrediting team or body have such methodological expertise.

In contrast, with the meaningful-processes approach, one need not worry overmuch about technical purity. Of course, one does not want to tolerate outright rubbish. But measures which provide *some* useful information are quite acceptable, provided the people involved in the assessment have a commitment to their use.

In this paper, we do not express a preference for one or the other of the two positions we have been discussing. At first blush, the final-answers approach has considerably more appeal. After all, if one asks whether an institution is achieving its objectives, presumably one would like an answer—a final answer— to the question and not simply find out who is involved in trying to answer the question, how they went about the task, etc. The meaningful-process approach seems solipsistic, a bottomless quagmire.

Despite the first-blush appeal of the final-answers approach, the experience of many people who have labored in this vineyard for a long time suggests that the meaningful-processes approach is the more fruitful one. The reasoning goes like this: Obtaining final, definitive answers to questions about outcomes is so difficult, so fraught with technical limitations, so long and arduous, that if you insist on final, clear answers, you are likely to despair; in frustration, you will give up the pursuit altogether. Further, since only the methodological experts were involved in the first place, no one much cares if the pursuit is abandoned. Better, so the reasoning goes, to concentrate on establishing the processes that will ensure continuing pursuit of answers to our questions about outcomes; and to involve the many people who have a stake in the answers and in the processes that can influence change in the answers.

It is an interesting contrast in positions. Those engaged in the accreditation business need to acknowledge the existence of these two different approaches and be ready to discuss them among themselves and with representatives of institutions served by accrediting agencies.

Program Effectiveness vs. Institutional Effectiveness. In this paper, we treat outcomes assessment related to *institutional* effectiveness. A companion paper is being written as part of the overall "COPA Outcomes Analysis Project" on outcomes assessment related to *program* effectiveness. What are the differences, if any, in the issues to be considered when treating programs vs. institutions? Is an institution's effectiveness defined as the simple arithmetic sum of the effectiveness of its various programs? If so, the problems of measuring institutional effectiveness are no more than the problems of measuring program effectiveness. To some extent, the answer to these questions depends on what is meant by "programs."

In most contexts, "programs" means academic programs designed by academic departments (faculty) or analogous bodies to train or develop students, particularly with respect to intellectual content or skills. In this context, institutional effectiveness (except for a few, very specialized institutions) goes well beyond the simple sum of measures for programs. Nearly all institutions specify desired outcomes beyond those related to their academic programs. Principal among such other outcomes are (a) development of non-academic outcomes for students such as values, attitudes, perspectives, etc., and (b) outcomes largely unrelated to students, such as community service, research productivity, etc. Institutional effectiveness, obviously, must be assessed for all such outcomes.

Of course, one could define "programs" very broadly, going well beyond the traditional definition of academic programs, to include all systematic efforts which an institution undertakes to accomplish its manifold goals. With this expanded definition, is there any difference between institutional effectiveness and the simple sum of the effectiveness of the institution's many programs? Yes, there is - or, at least, there may be.

It may be that the various programs sponsored by an institution *interact* in such a way that certain transcendent institutional goals are achieved (or fail to be achieved) beyond what can be accounted for by examining the successes and failures of individual programs. (If the statistically trained reader wonders if we use the term "interact" in the same sense in which it is used in analysis of variance, the answer is an unqualified "yes.") Hence, institutional effectiveness requires examination beyond simple tallying of the degree of effectiveness of all the separate programs of the institution.

The Importance of Goals. In the midst of much disagreement on many topics related to assessing the quality of educational institutions, there is one proposition on which everyone seems to agree: Quality, at least to the extent that it is defined by outcomes, *cannot* be evaluated or judged without reference to the mission, goals, or objectives of the institution. (There are some useful distinctions among these three terms—mission, goals, objectives—in certain contexts, but for our purposes we use them more or less interchangeably.)

Agreement on the proposition that outcomes assessment must start with institutional goals is a function, at least to some extent, of the diversity of higher education as it is practiced in the United States and similar countries. While all or nearly all of these institutions have some goals in common, chiefly the development of intellectual capacities, communication skills, and career preparation, there are also noteworthy differences among the institutions.

Many efforts to build an outcomes assessment plan for an institution prematurely consider the assessment methods to be employed. In fact, the first step in the process must be a careful analysis of the institution's goals. Methods, measures, indices, and procedures are then developed for specific application to these goals. Hence, well developed assessment plans always eventuate in some type of goals-by-methods matrix, such as that represented below. The task of those responsible for developing an assessment plan then becomes a matter of judiciously identifying which cells will be active in the matrix, i.e., which methods are appropriate for which goals. This idea of a goals-by-methods matrix should be kept in mind as we treat the various methods which are the main topic of this paper.

A Note on Reliability and Validity It is inevitable, as it should be, that when issues of methods of assessment or measurements are discussed, the questions of the reliability and validity of the measures will be raised. In our treatment of methods in the next section, we will *not* be discussing these issues. They are enormously important matters. However, they must be treated with respect to very specific uses of particular measures, especially with respect to validity. Hence, at this point, we simply note the importance of raising questions about validity and reliability of measures in the context of an institution's individual assessment plan; but we do not treat these matters in detail in the following discussion of methods of assessment.

	Methods, Measures			
Institutional Goals	A	B	C	etc.
1				
2				
3				
4				
etc.				

Methods for Assessing Outcomes

Considering that the main topic for this paper is "methods for outcomes assessment," the reader may, by now, be somewhat impatient to get on with the meat of the matter—which indeed we do take up in this section. However, we re-emphasize the need to reflect on the questions treated in earlier sections before treating particular methods and instruments, since directions set by the earlier considerations will influence the selection of methods and instruments.

As we begin the treatment of particular methods and instruments, two general observations should be made. First, the methods available for outcomes assessment are virtually infinite, being limited only by the types of outcomes specified for an institution and the imagination of those concerned with assessing the outcomes. There is *no* fixed list of methods. In what follows, we identify categories of the more widely used methods and types of instruments, but there is no suggestion that the list presented here is definitive and exhaustive.

Second, when presenting particular methods or types of instruments, we may occasionally refer to an example of a currently used instrument. But we have not attempted to produce a complete catalog of specific instruments, e.g. of all the currently used standardized tests aimed at assessing outcomes of general education programs. When we do refer to a particular example, it should not be inferred that this is an endorsement of the example as the best or most frequently used item within its general category.

Tests of Developed Abilities, Knowledge, and Skills

The first major category of methods for assessing outcomes for educational institutions consists of *tests*[2]. The word "tests" can be construed broadly to include any systematic source of information about characteristics of individuals (whether those "individuals" be people, institutions, or other entities), in which case "tests" becomes virtually synonymous with the word "methods" as used in this paper. However, the more popular use of the term "tests" refers to such things as intelligence tests, final examinations, the GRE's, etc. It is in this more popular sense that we use the term here.

At the heart of every educational institution is the intent to develop in students certain areas of knowledge, abilities, and/or skills. Ordinarily, the most direct assessment of whether or not these outcomes have been achieved is provided by students' performance on some kind of test. On that proposition there is near universal agreement. The real difficulty comes in answering the following three questions:

1. Does there exist a test (or group of tests) which validly assesses the particular outcomes of interest to an institution?
2. If not, is it possible to construct such a test (or group of tests)?
3. Assuming some degree of positive answer to the first two questions, how does one specify the level of acceptable performance on the test(s)? (This is the classical problem of standard-setting which has received so much attention in the accountability movement.)

Let us try to provide some guidance in answering these questions as they relate to measuring institutional effectiveness. It will be convenient to organize the matter in terms of the chart shown below.

		Areas to be Covered	
		General Education	Specific Areas
Source of Development	External		
	Local		

Areas to be Covered. Within the broad category of knowledge, abilities, and skills, we distinguish first between the areas to be covered, and subdivide the areas into general education and specific areas. The specific areas are usually represented by content fields corresponding to majors or programs, such as psychology, history, marketing, auto mechanics, etc. Each of these specific areas, in turn, can be resolved into ever-increasing levels of detail.

For some institutions, the general education category is simply irrelevant; the institution has no goals that can be categorized as general education. Such might be the case for institutions with an entirely vocational orientation; it will also be true for major schools (e.g. graduate schools) within institutions which concentrate only on specific areas of knowledge or skill. On the other hand, nearly all four-year colleges and universities consider general education outcomes to be among their most important goals.

All educational institutions have goals related to specific areas of knowledge, abilities, or skills.

Sources of Development. The other basis for categorization of tests is in terms of the source of development for the tests. Some are externally developed, i.e. outside of the institution. External developers are typically are of two types. First, there are "commercial" publishers (either profit or non-profit) whose business is, at least in part, the development of educational tests. Examples of such publishers are Educational Testing Service, the American College Test Program, the Psychological Corporation, Riverside Publishing (formerly part of Houghton-Mifflin Co.), and McGraw-Hill/ California Test Bureau. Second, there are professional associations which, as part of their service to their members, provide tests applicable to their particular areas of concern. Examples of such associations—there are hundreds of them—are the National League for Nursing, the American Chemical Society, and state Bar Associations.

The other source of development is local, i.e., people at the institution develop their own test. Such development may take place entirely within a single department; it may be done by an institution-wide committee; or it may be done as a combined effort of a department or committee and test experts in an educational research office at the institution.

We have not represented in our scheme for "sources of development" an intermediate category which might involve groups of institutions cooperating in development of a test. While this is certainly a theoretical possibility, the simple fact of the matter is that there is little of it taking place. (A notable exception is the development of the Area Concentration Achievement Tests, a FIPSE-funded project directed by Anthony Golden at Austin Peay State University.) People who are using tests to assess knowledge, abilities, and skills generally either use an externally developed test or develop one strictly within the confines of their own institution, although with some borrowing of experience from other institutions.

The strengths/advantages and weaknesses/ disadvantages of *tests* as methods for assessing institutional outcomes are partly the same and partly different for each of the cells in the chart given above. Before beginning to analyze these strengths and weaknesses, let us comment briefly on the nature of each one of the cells.

Externally Developed Tests of General Education. Externally developed tests aimed at the general education outcomes of post-secondary education are a relatively new phenomenon. They have started to become available as the "assessment movement" has grown in higher education in just the past fifteen years or so. Whereas broad-based measures of school achievement have been available at the elementary and secondary levels since the 1930's and such measures have long been used in nearly all elementary and secondary schools, comparable measures at the post-secondary level are still few in number and infrequently used.

However, there are now available several examples of measures (what are often called "batteries" of tests) developed by professional test-makers which are aimed at the general education component of a college degree program. Perhaps the most widely referenced is the COMP (College Outcomes Measurement Project) Test, produced by the American College Testing. Other examples include College BASE, produced by Riverside Publishing Company; the Academic Profile, produced by Educational Testing Service; and the Collegiate Assessment of Academic Proficiency (CAAP), produced by American College Testing.

One will sometimes hear reference to possible use of the Scholastic Aptitude Test (SAT) or the American College Test (ACT) with college seniors as a "post-test" following use of the same measures prior to admission as an attempt to assess growth in general education. But such a position is not accorded much credibility by measurement experts because these two widely used tests were specifically designed to be general predictors of success in college rather than as measures of the achievement of college outcomes. (One must admit, however, a considerable communality, from a strictly empirical viewpoint, in what is covered by tests designed for these differing purposes.)

Locally Developed Tests of General Education. There is, at present, a great deal of activity devoted to local development of tests related to general education goals. In some instances, institution-wide committees are attempting to develop "batteries" of tests along the lines of externally developed tests, but (presumably) more carefully attuned to local definitions of general education goals. In other instances, institutions are attempting to "embed" tests related to general education goals in existing course evaluation procedures, then aggregate results across courses.

Externally Developed Tests of Specific Areas. Externally developed tests of specific knowledge, abilities, and skills have been available for certain areas for many years. Examples include the advanced or subject tests of the Graduate Record Examinations, the Major Field Achievement Tests, both of the latter being published by the College Board, and the tests administered by professional organizations, such as the National League for Nursing (NLN) tests and CPA examinations. More such tests become available with each passing year.

Locally Developed Tests of Specific Areas. Locally developed tests of specific knowledge, abilities, and skills are legion. They include comprehensive, end-of-program examinations in every conceivable content area. They are typically developed by the faculty responsible for a particular program. They may look very much like an externally developed exam such as a GRE or they may be combinations of written tests and behavioral ratings of skill performance.

A Special Note on Portfolios. Much work is being carried out these days on the use of "portfolios" for assessment of student learning. The work is generally being conducted by local institutions; some of it is aimed at general education goals, especially with respect to writing and speaking skills, while in other instances it is aimed at major fields of study. Portfolio assessment offers some interesting possibilities, particularly in contrast to what are referred to as traditional paper-and-pencil tests.

In terms of the questions to be asked about this particular approach to assessment and analysis of its strengths and weaknesses, portfolio assessment is not generically different from other types of tests. That is, one must be concerned about questions of validity, reliability, efficiency, interpretability, and so forth, for this source of information about achievement in the same manner as considering these matters for other sources of information.

General Strengths and Weaknesses of Tests. As noted above, strengths and weaknesses of tests as assessments of outcomes must be analyzed, at least to some extent, separately for each cell in our scheme, especially with respect to the source of development. However, there are some generic strengths and weaknesses of tests which we can discuss first.

The major strength of tests as an outcome assessment is very obvious. At the heart of most educational enterprises is the development of knowledge, ability, and skill. A test, at least a well developed one, is the most direct way to "get at" such development. A test has a directness to it which is not present in most other methodologies for purposes of assessing development of knowledge, ability, and skill. There is good reason why tests play such a central role in most educational settings.

A second strength of tests is that they have fairly wide applicability—partly because of the way most instructional goals are conceptualized and the way instruction is delivered. The way we educate students usually lends itself to testing the outcomes. Educational critics sometimes lament this situation, but the simple fact is that the way in which much education takes place lends itself to what might be called traditional testing methodology.

A third strength of tests is that we now have a fairly well-developed set of principles, contained within the field of psychometrics, for determining, or at least discussing, the quality of measurements provided by tests. These principles are much more developed than they are for the other methods to be considered later in this paper.

Tests have three major weaknesses for purposes of outcomes assessment related to institutional effectiveness. First, tests are appropriate for only a few of an institution's goals, although these few goals are usually among the most important for the institution. This is not a fault of tests as such but a fault in our typical thinking about outcomes assessment. We alluded to this phenomenon earlier. When asked about assessing outcomes for an institution, a frequent initial response is: There isn't a test for all the things we are trying to do. Indeed, that may be true, and almost certainly is true. But there are tests for some of the goals, and other methods for other goals.

The second weakness of tests as a generic methodology is the fact that we usually cannot combine information across many different fields. This weakness is peculiar to assessment of *institutional* effectiveness as opposed to program effectiveness. The typical institution has many programs. Some of these programs may have well developed tests applicable to their particular outcomes; other programs may have poorly developed tests or no tests for their outcomes. Usually, there is no meaningful way to combine all of this information (and lack of information) to make statements about the institution as a whole. Even if every program had a "good test" for its outcomes, it is unlikely that the information could be combined in any meaningful way. One program may use the GRE's, another the NLN exams, another a locally developed paper-and-pencil test, still another a laboratory-based practicum as an exit measure. Each of these may provide a good measure for an individual program, but generalizing across the entire institution is difficult at best, demanding a level of inference that could only be described as ethereal.

A third generic criticism of tests relates to the relationship between test performance and actual practice or application, usually defined in a post-graduation time frame. People always wonder whether the knowledge displayed on a test will translate into proper application later on (or even whether the knowledge will be retained six months later).

Strengths/Advantages of Externally Developed Tests. With the understanding that the strengths and weaknesses discussed above apply to the following discussion, let us turn now to specific categories of tests. Externally developed tests have a number of strengths or advantages, most of which (but not all) apply to both tests of general education and tests of specific areas.

The first advantage of externally developed tests is the professional care taken in their preparation. Ordinarily, the developers of the tests have a high degree of expertise and have devoted a considerable amount of time and research to the test development process.

The second advantage of externally developed tests is that (usually) they have some type of norm, i.e. an external reference point for comparing local performance. This is a major advantage for "making sense" out of local findings.

The third advantage of externally developed tests—a fairly obvious one—is the ready availability of materials and services. If an institution decides that a particular test is appropriate for some use, it can usually order the materials and have them on hand in a matter of weeks. Further, such services as machine scoring and reporting, consulting, technical manuals, etc. are already in place.

A fourth advantage of externally developed tests is their credibility with a variety of audiences. Whether the credibility is justified or not is a separate matter. The fact is that many groups "pay attention" to GRE scores, passing rates on the CPA exams, etc.

A final characteristic of at least some externally developed exams, which should probably be listed as an advantage or strength, is that they serve as *the* definition of successful performance in the field. There is simply no way to avoid them, other than dropping out of the field. This characteristic applies in many "professional" fields such as nursing, physical therapy, and education.

Weaknesses/Disadvantages of Externally Developed Tests. The greatest disadvantage of externally developed tests is, quite simply, that they do not exist for many fields of knowledge or skill which represent important outcomes within an institution.

A second disadvantage of externally developed tests, really a milder form of the first disadvantage, is that the test may only partially cover the outcomes specified at the local institution. This is, of course, a matter of degree. And it is the source of much controversy when contemplating the use of an externally developed test. How much overlap does there need to be between locally specified outcomes and the content of an externally developed test before one declares that the "fit" is close enough?

A third disadvantage of externally developed tests, very much related to the latter point, is their relative inflexibility. In most cases, using such a test is an all-or-none affair. One cannot ordinarily eliminate some of the content and/or add locally developed content, although some externally developed tests give limited flexibility in this regard.

A fourth (potential) disadvantage of externally developed tests is entirely psychological in nature. It often afflicts faculty involved in specifying what measures are appropriate for their program. There is, sometimes, not a "sense of ownership" when an externally developed test is used. Because the test cannot, in most instances, be tailored to the local curriculum— even a little bit—there is a tendency to reject the whole notion of using an external test. Of course, not all faculty feel this way; some welcome the objectivity of an external measure. But the lack of "ownership" feeling is encountered sufficiently often to warrant mention here. A related concern is that "tests will drive the curriculum" or that faculty will "teach to the tests." These fears are mentioned more often with respect to externally developed tests, although they may be just as problematic for local tests.

A fifth disadvantage of externally developed tests is their costs. Because a great deal of expertise, research, and time has been devoted to their development and their developers would like, at a minimum, to not lose money, the costs are naturally passed along to the users.

With respect to this disadvantage, two comments are worth making. First, where cost is a real consideration, the cost can be reduced by sampling, either by student within year, e.g. testing a 20% sample of the students, or by student cohorts across years, e.g. testing students every third year. For purposes of *institutional* assessment, such sampling schemes are perfectly acceptable. Second, if costs of externally developed tests are compared with the *true* costs of local development of tests, it may actually be cheaper to use the external tests; but institutions rarely calculate the true costs of local development, being particularly adept at disregarding the huge amounts of faculty and staff time which may be devoted to the test development enterprise.

Strengths and Weaknesses of Locally Developed Tests. For the most part, the strengths of locally developed tests are the flip side of the weaknesses of externally developed tests; and the weaknesses of locally developed tests are the flip side of the strengths of externally developed tests. And, it is not difficult to see what the "flip sides" are. Let us treat the matter in summary fashion.

The strengths/advantages of locally developed tests are:

1. They are potentially available for *all* areas, dependent only on the local will to develop them.
2. They can be tailored *exactly* to fit the local definition of outcomes.
3. They are infinitely flexible. They can be modified from year to year or even for various sub-groups of students.
4. The sense of ownership can be very strong for locally developed tests. This is perhaps their major advantage. Rather incidentally—but very importantly, especially if one opts for the "process" approach to assessment—the engagement of local personnel in building a test usually leads to intense discussion of what the goals of education really are, how they can be achieved, what changes can be made to achieve them, etc.—all of which might be far more important for the overall quality of the institution than any results of the test.
5. It can be relatively inexpensive to develop tests locally (at least if one disregards the personnel costs involved).

The weaknesses/disadvantages of locally developed tests include:

1. Expertise in test development may be in short supply, with attendant deleterious consequences. It is hard to build a good test; it is easy to build a poor one. In particular, locally developed tests tend to either one of two extremes: either very factual and mundane or so esoteric as to be uninterpretable.
2. Locally developed tests almost never have any external reference points (norms). When confronted with the final results of the test, it is almost impossible to answer the question: Is this result good? It may be possible to have faculty, with a thorough knowledge of the local curriculum, establish some "criterion-referenced" benchmarks. But this is often a more difficult process than one first supposes it to be. Of course, after a locally developed test has been used more than once, it will have its own "local norm" which can be used as a reference point for future uses.
3. It is a long and arduous journey to develop a local test. From the point of deciding to develop a test to the point of having a test which is useable is often several years—and sometimes you never get there. Quite apart from developing the test itself, one must worry about procedures for scoring them, summarizing results, printing the materials, handling inquiries about technical characteristics of the test, etc.
4. Locally developed tests rarely have the credibility with outside audiences (i.e., outside the faculty/staff who developed them) that external tests have.
5. The fifth item listed as a "strength" of externally developed tests (i.e., being the "coin of the realm" for certain agencies) does not really have a corresponding weakness for locally developed tests, except the perfectly obvious one that local tests do not serve this function.

Surveys, Questionnaires

The second major method employed for assessing educational outcomes is the very broad category of surveys and questionnaires. For economy, we will use the single term "survey" in this paper to represent this category. The category includes such a diffuse, flexible, varied number of applications and examples that it is, at first, difficult to summarize. In particular, the content—what you ask questions about—in surveys is limitless. Nonetheless, let us attempt to summarize the category with the following organizational scheme: by source of development, target group, format, and content orientation.

Sources of development. The sources of development for surveys are the same as for "tests," with one or two exceptions. First, there are a great variety of externally developed surveys. They are available from commercial publishers, an example being the Alumni Survey published by the American College Testing Program. They are also available from professional associations, for example the survey sent to graduates of rehabilitation programs by the Council on Rehabilitation Education when rehabilitation programs are being accredited. Several well-known surveys are also available from educational research organizations, which function, in this regard, partly like commercial publishers. Two examples in this category are the College Student Experiences Questionnaire (sometimes called the Pace survey, after its principal author, Robert Pace) produced by the Center for the Study of Evaluation at UCLA; and the Freshman Survey (sometimes called the Astin survey after its principal author, Alexander Astin) produced by the Cooperative Institutional Research Program also at UCLA.

Then, of course, there are locally developed surveys, by the hundreds, even thousands. Every institution has a great variety of locally developed surveys, many of which relate, in one way or another, to outcomes assessment.

There are also a number of examples of groups of institutions cooperating in the development and use of surveys for outcomes assessment, for example for follow-up of graduates. Under sources of development for tests, it was noted that institutions don't seem to cooperate in developing tests; but they do cooperate in developing survey instruments.

Target Groups. Target groups for surveys are virtually limitless, but there are about a half-dozen typical target groups, including the following:

1. Students, further subdivided by level, e.g. pre-admission level, entry level, in-process (e.g. end of sophomore year), and exit level (approximately the time of graduation).
2. Graduates, further subdivided by number of years after graduation, e.g. after six months, one year, five years, ten years, etc.
3. Employers of graduates of the institution.
4. Faculty and staff of the institution, obviously susceptible to any one of several further subdivisions, e.g. humanities faculty, junior vs. senior faculty, academic vs. non-academic administrators, etc.
5. Community members, again obviously susceptible to a variety of subdivisions, e.g. community leaders, a random sample of area residents, etc.
6. Personnel at peer institutions.

Format. It should be mentioned, perhaps as part of the definition of what is meant by "surveys," that this method encompasses a number of possible formats. The most typical is a printed form, on which the respondent—the member of the target group—records answers to items. The items may be of the "check off" variety using rating scales, yes-no answers, etc. and/or of the "free response" variety in which the respondent writes out an answer to a question.

Since most people think of surveys as asking only about subjective feelings, special mention should be made here of the "behavior checklist" type of item which may be used in a survey. In this type of item, respondents are not asked their opinions about something but whether or not they have performed some action or how often they do it. For example, students may be asked how often they visited the library last week. Faculty may be asked how many students visited their office last week. Alumni may be asked how many professional associations they belong to.

We also include within the scope of surveys any type of interview. These may take place in person or by telephone. They may also occur within the context of a "focus group" which essentially involves interviewing a small group of persons all at once. Any of these formats for interviews may vary from highly structured, almost like a "check off" questionnaire, to quite unstructured affairs.

Content. It is difficult to characterize the possible content of surveys. This difficulty is precisely the strength of surveys. One can construct a survey—or find an existing one—to cover just about anything. However, we can list some of the types of items often encountered in surveys which are relevant to the question of outcomes assessment for institutional assessment, while re-emphasizing that one can cover almost anything with survey methodology.

Surveys of students often ask about three main topics. First, after some period of association with the institution, students are asked to rate their perceived growth with respect to knowledge, skills, or abilities. Second, students are queried about their attitudes, dispositions, and habits, for example with respect to matters related to life-long goals, racial and social attitudes, religious and civic practices, etc. Third, students are asked to rate various aspects of the institution, from the mundane (food service, parking facilities, etc.) to the sublime (overall satisfaction with instruction, opportunities for research, etc.).

Surveys of graduates cover much the same territory as surveys of students, with less emphasis on the more mundane characteristics of student life, which may have changed substantially (for better or worse) since the graduate was in attendance. Additionally, surveys of graduates inquire about current job status, relationship between job and degree program, and involvement in community and professional activities.

Surveys of employers usually attempt to determine whether the employee who is a graduate of the institution possesses certain required skills and whether the employer is satisfied with the employee/graduate.

Surveys of community members usually relate to those institutional goals which are oriented to community service. Do community members perceive the institution as providing important community services, e.g. cultural events, assistance to local businesses, etc.?

Surveys of faculty and staff usually ask for ratings of the extent to which these employees see the institution as achieving its goals. Such surveys often add questions about employee satisfaction with various institutional services (much like the questions asked of students) and about satisfaction with working conditions, questions in these two latter categories being useful for purposes of human resources management but perhaps not of much relevance for outcomes assessment.

In this discussion of content orientation, special mention should be made of instruments which were originally designed to be personality tests, but are occasionally used for purposes of outcomes assessment. Examples include the Omnibus Personality Inventory and the Allport-Vernon-Lindsey Study of Values. Although originally designed for purposes other than educational outcomes assessment and although technically "tests" hence includable within the earlier section of this paper on various types of tests, these personality tests seem to fit best, functionally, in our treatment of surveys in this paper.

We emphasize again that the above listing simply provides examples of the many topics which can be covered by surveys and of the types of target groups for these surveys.

Strengths/Advantages of Surveys. As we begin the analysis of strengths and weaknesses of surveys for use in outcomes assessment, we should note that in terms of the contrast between externally developed and locally developed surveys, the strengths and weaknesses are almost identical to those discussed with respect to this issue under "tests" above. To summarize just briefly, strengths of the externally developed instrument (test or survey) are immediate availability of materials and services, quality of development, external reference points (norms), and credibility to outside audiences. The weaknesses tend to be lack of fit with local emphases and circumstances, political acceptance (ownership), and possibly costs.

There is one exception to this generalization. Externally developed surveys often, although not universally, allow for inclusion of locally developed questions (items) as a supplement to the main part of the survey. This tends not to be true for externally developed tests. Hence, the criticism that externally developed tests have little flexibility for adaptation to local emphases must be tempered, although not eliminated entirely, as applied to externally developed surveys. (The real difference here is not so much between surveys and tests as between the typical structure of surveys and tests, particularly as that structure serves as the basis for interpretation. Many surveys are interpreted on an item-by-item basis, thus adding items, which will also be interpreted on an item-by-item basis, is not problematic. On the other hand, many tests are interpreted in terms of scores derived from a summation of responses across items; and it is at least messy, often impossible, to insert extra items which will be summed into the total scores.)

The overwhelming strengths of the survey as a method of assessing outcomes are its exceptional flexibility and near universal applicability. One can construct or find an existing survey to cover virtually any topic, for any type of target group. This makes the survey a particularly attractive alternative for "getting at" outcomes which do not seem to be susceptible to other types of assessments. One can always just ask students, faculty, community members, etc., whether they believe that some outcome has been achieved, so that at least *some* information is available on the matter.

Part of the flexibility of surveys relates to their administration. Ordinarily, the administration does not need to be tightly controlled. Surveys can be mailed so that respondents can complete them on their own. Or they can be administered in group settings, without worrying overmuch about security, cheating, etc. This is in contrast to cognitive tests for which administration often needs to be tightly controlled.

Another advantage of the survey is that it can be relatively easy to construct. It certainly takes some degree of expertise to do a reasonable job of constructing a survey (or local, supplementary items to be used with an externally developed survey). But the degree of expertise required tends to be considerably less than what is normally required to construct a good cognitive test.

There is one advantage of surveys that is specially relevant to their use for institution-wide assessment as opposed to program assessment. Survey questions can be phrased in such a way that they apply to whatever program the respondent is associated with (for example, students in their various majors, faculty in their various departments), with responses summarized across all programs, i.e., across the entire institution. This contrasts with the difficulty of combining *test* information across various programs, as noted in the discussion of tests.

A final advantage of the survey—at least in most instances— is that persons who are not methodological experts usually feel relatively comfortable looking at survey data. Such data is usually reported in the form of percentages of responses to individual items, hence is fairly simple to handle. This is in contrast to reports on standardized tests such as the GRE's or MCAT's where scores are given in "standard scores" or other modes unfamiliar to the non-expert. This is an important advantage for institution-wide assessment, which usually involves audiences (including accrediting teams and associations) which are diverse with respect to their interests and levels of methodological expertise.

Weaknesses/Disadvantages of Surveys. There is one overwhelming weakness to the survey for purposes of use in outcomes assessment (or for just about any other purpose): The responses are generally *subjective* assessments. Thus, the student may say, on a survey, that he has grown tremendously in writing ability as a result of his college experience, when in fact he has gotten no better at all. A student may say, on a survey, that the university library is horrible, when in fact, according to a number of objective criteria, it is quite a good library. Graduates may say, on a survey, that they are very tolerant of persons from other races and cultures, when in fact, any objective analysis of the graduates' behavior contradicts this report. A faculty member may say, on a survey, that pitifully few students are involved in research projects at the institution, when in fact the fifteen percent of students who are involved is a relatively high percentage for institutions of this type. And so it goes with other types of responses to surveys. They can provide us with a great deal of information but we must always be asking about the correspondence between reality and the subjective responses given on the surveys.

The situation is, of course, not hopeless. Much research has been done on the validity of survey responses; and much of that research suggests that many surveys have some degree of validity. However, the correspondence between reality and survey responses is far from perfect and, worse, we are never quite sure when the correspondence might be nil. Hence, we must always be looking for ways in which to supplement survey data.

A final weakness of surveys is the flip side of the strength we identified above as ease of construction. While it is *relatively* easy to construct reasonably good surveys, it is also easy to construct some really horrible ones. And although most people recognize that some expertise is required to construct cognitive tests, there are too many people who believe that any neophyte can "throw together" a survey: A dangerous attitude. Care must be exercised in survey construction if meaningful results are to be obtained.

Institutional Records

The third major category of methods which may be used for the assessment of outcomes for institutions is that of institutional records. This is probably the most neglected of all types of methods for assessing outcomes. It seems to be a matter of not seeing the forest because of all the trees.

Every institution keeps a great variety of records. Nearly all of them are kept for some reason other than specifically for the purposes of assessing outcomes. Nevertheless, many of these records can, in fact, be used for outcomes assessment. Often, all that is required is a little imagination—and finding out who keeps what records and getting their cooperation in using the records for something other than their primarily intended purpose. We provide here a number of examples of the use of institutional records for outcomes assessment, prefacing each example with reference to a *possible* institutional goal.

1. An institution which aims to serve minority students in its region can use information from its admissions database to answer many questions. How many "inquiries" have been generated from potential minority students? How many applications? How many actual first-time registrations? How have these figures changed over the years?
2. An institution which aims to have a strong liberal arts component in its bachelor's degree programs can analyze

transcripts of graduating seniors to determine if, in fact, students have pursued programs which can be characterized as liberal arts in orientation. (It should be noted that in many instances, institutions have curricular regulations which seem to require a liberal arts orientation, but students find ingenious ways to circumvent the regulations.)

3. An institution which aims to encourage research on the part of its students can determine, from institutional records, such things as: (a) numbers of published papers or presentations at professional meetings which involve joint faculty-student authorship, (b) number of grant proposals which include money for student assistants, (c) patterns of library usage by students, e.g. number of books checked out.
4. An institution which aims to foster community involvement on the part of its students can track patterns over the years in the numbers of students volunteering for various community projects, number of student organizations involved, numbers of hours devoted to the projects, etc. An institution may very well have one office coordinating volunteer activities and this office must have this type of information just to serve its coordinating function.
5. An institution which wants to encourage further study by its students can track the number going on to higher levels of education, e.g. to a four-year college from a community college, to doctoral programs following a bachelor's degree program. Such information is often obtained as a by-product of job placement data collected by career service offices.
6. An institution which wishes to encourage participation in religious activities can note attendance patterns for chapel, mass, religious retreats and other such functions.
7. An institution which purports to serve the local business community can report the number of businesses served through agencies such as a Small Business Development Center, a technology center, or other outreach offices.

Strengths/Advantages of Institutional Records. For purposes of use in outcomes assessment, institutional records have a number of very significant strengths or advantages. First, and perhaps foremost, the information has already been collected, hence a minimum of additional effort is required—although some effort is needed, as noted below. Second and relatedly, the information has often been collected over a period of time, perhaps for many years, thus allowing for identification of trends. And, it is likely that whoever collects the information will continue to do so in the future. Third, the information often has a directness and simplicity to it, a kind of "face validity," which is refreshing. Finally, by their nature, institutional records ordinarily do relate to the entire institution rather than to just a single program or small segment of the institutional community.

Weaknesses/Disadvantages of Institutional Records. Institutional records also have certain weaknesses or disadvantages with respect to their use for purposes of outcomes assessment. First, while the records do exist, it is often the case that someone whose principal concern is outcomes assessment must collate, summarize, analyze, and interpret the information contained in the records; the records do not often speak for themselves. Second, since the records are often created originally for some purposes other than those of primary concern in this paper, one often wishes that the records had been created and stored in a somewhat different manner: something which one can only lament after the fact. In relation to this point, it often happens that the person, group, or office interested in using the institutional records (which are under the control of some other person, group, or office) for purposes of outcomes assessment wishes to change the way the records are created or stored. Suggestions to do so, experience tells us, must be done with the utmost tact. The principal custodians of the records usually do not like "tampering" with their systems; they are usually suspicious of any "outsiders" who want to use their data and suspicion can turn to outright hostility when the outsider not only wants to use the information but actually change how it is obtained.

A final weakness—perhaps better termed a caution—to be noted about the use of institutional records for outcomes assessment is that one must become thoroughly familiar with the way in which information is collected before relying on it heavily. The simplicity of the records can sometimes be deceptive. The operational definition of some count may be quite different than the common-sense understanding of the name given to it. For example, a report of "numbers of students attending a cultural event" may be based on an actual "turnstile" count, or it may be based on someone's "eyeball" estimate of the size of the crowd. "Number of applications from minority students" may mean number of application forms received (regardless of whether all necessary supporting materials are received) or it may mean number of complete application folders processed (with all necessary supporting documents). "Number of businesses served" may include even casual phone inquiries from any business in the area, or it may mean at least one formal meeting with the business followed up by some type of written report.

The offices which originally collect such information know the details of these matters and they usually have good reasons for collecting the information in a particular manner. Persons who now want to use the information for purposes of institutional outcomes assessment must become thoroughly familiar with the way the information is collected and tailor conclusions accordingly. If the information is used in comparisons with other institutions, it is obviously important that all institutions involved are using the same operational definitions.

It is also important that persons involved in the accreditation enterprise be sensitive to this matter of the operational definitions used for institutional records. When such records are used to demonstrate institutional effectiveness in the attainment of some goal, it is sometimes necessary to do a little probing in order to understand exactly what the records are showing.

Concluding Note on Institutional Records. Institutional records are perhaps the most underutilized of all methods for purposes of outcomes assessment. All institutions collect great amounts of information. Ordinarily, the information is collected by one office for its own purposes, without much thought about its use in some wider context of institutional assessment. Persons responsible for thinking about institutional assessment should devote some time to determining what information is routinely collected by various offices in the institution, then reflecting—or, perhaps better, brainstorming—about what parts of it might be used in an overall institutional assessment plan. Personnel involved in the accreditation process should be alert to look for institutional records which may reveal important information about achievement of institutional goals, even when the information is not presented as part of the institution's "assessment" information.

The Institutional Assessment Plan

Following some introductory comments on general issues which must be taken into account when considering methods of assessment, we have concentrated on identification of particular methods for outcomes assessment and their strengths and weaknesses, especially with respect to institutional effectiveness. In practice, all of these thoughts and information must be combined into some type of overall plan for the assessment of outcomes. Constructing such a plan is a large task. It is a topic not treated in this paper. However, we want to conclude the paper by noting the importance of the overall plan. It is in this plan that all of the issues covered in this paper are treated "live." Most frequently the plan will incorporate some mixture of externally developed tests, locally developed tests, externally developed surveys, locally developed surveys, and certain institutional records. Some parts of the plan will be well developed, others poorly developed, at least for a time. Again, it is not the purpose of this paper to treat the construction and implementation of such plans in detail, but our consideration of methods would not be complete without reference to the fact that the methods are ultimately incorporated into some type of overall assessment plan.

*Thomas P. Hogan is Dean of the Graduate School at the University of Scranton.

Endnotes

1. Throughout this paper, we have avoided citations in the professional literature. However, an exception is made here to cite two works which relate directly to the methodological difficulties involved in using the value added approach. The two works, both recently published and both being chapter-long appendices to books, provide excellent summaries of these problems and attempts to deal with them. The first is an appendix entitled "Statistical Analysis of Longitudinal Data" in Astin, A.W. (1991). *Assessment for Excellence: The Philosophy and Practice of Assessment and Evaluation in Higher Education*. New York: American Council on Education/Macmillan Publishing Company. The second is an appendix entitled "Methodological Issues in Assessing the Influence of College" in Pascarella, E.T. and Terenzini, P.T. (1991). *How College Affects Students*. San Francisco: Jossey-Bass Publishers.

2. In listing tests first in order of presentation here, we do not mean to imply that they have some type of qualitative primacy, i.e., that they are the best or most important types of assessment. The order of presentation is purely arbitrary. What is the best or most important type of assessment depends on the goal that is being addressed.

Selected References on Outcomes Assessment in Higher Education: An Annotated Bibliography

Trudy W. Banta*

Foreword

This bibliography on postsecondary outcomes assessment is by no means exhaustive. As the title of the work implies, the several individuals and groups who provided guidance for its development made conscious selections from the rather substantial body of literature that has grown up since 1985 around the topic of outcomes assessment in colleges and universities. These selections were made in response to the question, "What are the key references that will give accrediting agency staff and those with whom they consult a comprehensive overview of the current status of the field of postsecondary student outcomes assessment and assessment of institutional effectiveness?"

While the bibliography is subdivided into a dozen sections for ease of reference, in some cases, the placement of a given work is somewhat arbitrary. Several of the citations could be placed in two or more categories, and the entries in the **Books/Collections/Review Articles** section contain material that belongs under several or all of the preceding sub-headings.

Though I must accept responsibility for the final decisions about materials to include or exclude, I would like to acknowledge the capable assistance I received in the process of developing the annotations from the following members of the staff at the Center for Assessment Research and Development: Margery Bensey, John Stuhl, Francine Reynolds, Gary Pike, and Ann-Marie Pitts.

Assessment and Accreditation

Commission on Colleges. Southern Association of Colleges and Schools. (1989). *Resource manual on institutional effectiveness. Second edition.* Decatur, GA: Author. This manual was designed to provide assistance on addressing the SACS institutional effectiveness criterion for those who conduct the self-study required for reaffirmation of institutional accreditation. The contents emphasize that self-study should be viewed not as an episodic event but as a continuous process aimed at improving the college or university. Assessment of institutional effectiveness is defined as a systematic comparison of institutional performance to institutional purpose.

Commission on Colleges. Southern Association of Colleges and Schools. (1991). *Program for peer evaluators: Training manual.* Decatur, GA: Author. This manual was developed to train members of accreditation visiting committees. The topics covered include: The peer review process, preparing for the visit, gathering evidence during the visit, applying specific accreditation guidelines in evaluating evidence, and preparing a report for the institution.

Commission on Institutions of Higher Education. North Central Association of Colleges and Schools. (1991). *Assessment workbook.* Chicago, IL: Author. This resource manual was developed for use in a series of regional seminars conducted in Spring 1991 by the North Central Association entitled, "Assessing Student Academic Achievement in the Context of the Criteria for Accreditation." Topics in the manual include: Implementing the NCA accreditation statement on assessment, developing an assessment program, and six case studies on institutional assessment of student academic achievement.

Folger, J.K., & Harris, J.W. (1989). *Assessment In Accreditation.* Decatur, GA: Southern Association of Colleges and Schools. This book is designed to give direction for developing an institutional assessment system that will provide ongoing information about effectiveness. It is aimed more particularly at institutions responding to new accreditation requirements for systematic and ongoing assessment of results. Appendices identify potential assessment instruments and other resources.

Ewell, P.T., & Lisensky, R.P. (1988). *Assessing institutional effectiveness: Redirecting the self-study process.* Washington, D.C.: Consortium for the Advancement of Private Higher Education. This book uses the results of a national demonstration project on the linkage between institutional assessment and regional accreditation to develop chapters on institutional goal definition, assessment for institutional distinctiveness, assessing general education, and organizing for assessment. An appendix provides an "Analytical Table of Contents for Self-Study" and formats for departmental data collection in connection with self-study.

Measurement Issues

Adelman, Clifford (Ed.). (1988). *Performance and judgment: Essays on principles and practice in the assessment of college student learning.* Washington, D.C.: Office of Educational Research and Improvement. A collection of essays including, Assessing the generic outcomes of higher education, Baird, L.L.; Assessment of basic skills in mathematics, Appelbaum, M.I.; Assessing general education, Centra, J.; Assessment through the major, Appelbaum, M.I.; Assessing changes in student values, Grandy, J.; Computer-based testing: Contributions of new technology, Grandy, J.; States of art in the science of writing and other performance assessments, Dunbar, S.; and Using the assessment center method of measure life competencies, Byham, W.C.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing.* Washington, D.C.: American Psychological Association. The fifth edition of the American Psychological Association's *Technical Recommendations for Psychological Tests and Diagnostic Techniques* addresses major current uses of tests, technical issues related to social and legal concerns and needs of all participating in the testing process. The *Standards* comprise four components: Part I, Technical Standards for Test Construction and Evaluation; Part II, Professional Standards for Test Use; Part III, Standards for Particular Applications; and Part IV, Standards for Administration Procedures.

Anrig, G.R. (1986). *A message for governors and state legislators: The minimum competency approach can be bad for the health of higher education.* Unpubished address. Princeton, New Jersey: Educational Testing Service. The president of the Educational Testing Service argues that tests measuring minimum competency are unfit for use in assessing higher education. Instead of these minimum competency measures, the author urges faculty at each institution to identify types of knowledge and particular skills they intend students to acquire, then to develop instruments designed to assess these learned abilities.

Baird L.L. (1988). A map of postsecondary assessment. *Research in Higher Education, 28,* 99-115. Students' knowledge and skills cannot be appropriately assessed in the

absence of knowledge about how their development is influenced by other aspects of postsecondary education. These aspects are described in a "map", consisting of twenty points, which depicts the flow of students through institutions and experiences from precollege to adulthood. The map suggests where better assessments and models are needed; for example, for adult learners, graduate and professional education, and plans of college seniors.

Banta, T.W., & Pike, G.R. (1989). Methods for comparing outcomes assessment instruments. *Research in Higher Education*, 30, 455-469. A general process is outlined for faculty use in comparing the relative efficacy of college outcomes assessment instruments for gauging student progress toward goals considered important by the faculty. Analysis of two standardized general education exams—the ACT COMP and the ETS Academic Profile—illustrates the process.

Baratz-Snowden, J. (1990). The NBPTS begins its research and development program. *Educational Researcher, 19*(6), 19-24. The author, vice-president for assessment and research at the National Board for Professional Teaching Standards, discusses the board's guidelines for setting nationwide standards for assessing entry-level skills of teaching candidates.

Council for the Advancement of Standards (CAS) for Student Services/Development Programs. (1988). *CAS standards and guidelines for student services/development programs: General-division level self-assessment guide*. College Park: The University of Maryland, Office of Student Affairs. Self-assessment guides focus on seventeen areas of self-study, such as admissions, orientation, career planning and placement, and counseling services.

Mines, R.A. (1985). Measurement issues in evaluating student development programs. *Journal of College Student Personnel, 26*, 101-106. An introduction to psychometric issues related to developmental assessment and recommendations for needed instrument refinements are provided. In addition, a brief overview of instruments available to assess developmental stages and tasks is included.

Osterlind, S.J. (1989). *Constructing test items*. Boston: Kluwer Academic Publishers. The author addresses the issues of functions and characteristics of test items, item formats, methods for assessing quality of test items, and issues related to use of test items. Chapter titles include: What is constructing test items?; Definition, purpose, and characteristics of items; Determining the content for items: Validity; Starting to write items; practical considerations; Style, editorial, and publication guidelines for items in the multiple-choice format; Style,editorial, and publication guidelines for items in other common formats; Judging the quality of test items: item analysis; and Ethical, legal considerations, and final remarks for item writers.

Pike, G.R., Phillippi, R.H., Banta, T.W., Bensey, M.W., Milbourne, C.C. & Columbus, P.J. (1991). *Freshman-to-senior gains at the University of Tennessee, Knoxville*. Knoxville: The University of Tennessee, Center for Assessment Research and Development. In 1989, a comprehensive study of student growth in college, as measured by freshman-to-senior gain scores on the College Outcome Measures Program (COMP) Objective Test was undertaken at the University of Tennessee, Knoxville. Analyses of scores for 942 students revealed the following: 1) students with the highest gain scores had the lowest entering aptitude and achievement scores; 2) gain scores were

not significantly related to measures of student involvement; 3) gain scores were not meaningfully related to coursework; and 4) serious problems existed with the reliability and validity of gain scores. Analysis of follow-up interviews with thirty students revealed that involvement and coursework were related to perceived growth and development.

Rogers, G. (1988). *Validating college outcomes with institutionally developed instruments: Issues in maximizing contextual validity.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA. The author clarifies the significance of contextual validity in the creation and use of assessment instruments. Using examples from the Alverno College Office of Research and Evaluation, Rogers demonstrates the elements and strategies practiced in improving the contextual validity of Alverno's assessment instruments.

Widgor, A.K., & Garner, W.R. (Eds.). (1982). *Ability testing: Uses, consequences, and controversies. Part I.* Washington: National Academic Press. This report is the work of the Committee on Ability Testing, functioning under the direction of the National Research Council. Its purposes are to describe the theory and practice of testing, to clarify the competing interests that bear upon the issue of testing, and to encourage improved, informed decision-making with regard to assessment. Part I presents discussions of testing issues, grouped in the following categories: Chapters 1-3 provide an overview of testing controversies, concept introduction, methods, terminology, a brief history of testing in the U.S., and legal requirements that have arisen around testing; Chapters 4-6 describe educational and employment uses of tests, common misuses, and recommendations for improved use; and Chapter 7 examines limitations of standardized tests and identifies contexts, i.e. cultural, societal, by which these limitations are affected.

State Policy Issues in Assessment

Aper, J.P., Culver, S.M., & Hinkle, D.E. (1990). Coming to terms with the accountability versus improvement debate in assessment. *Higher Education, 20,* 471-483. While many postsecondary institutions consider a mandate for outcomes assessment an accountability issue, state governments appear to be more concerned about improving higher education than about making summative judgments. These two perspectives frame a fundamental debate about the nature and direction of outcomes assessment.

Berdahl, R.D., Peterson, M.W., Studds, S., & Mets, L.A. (1987). *The state's role and impact in improving quality in undergraduate education: A perspective and framework.* College Park, MD: National Center for Postsecondary Governance and Finance. This study explores the role of state government in improving the quality of undergraduate education and the impact of related state actions on colleges. Following a review of the background for state interest in quality, state approaches to quality improvement are considered, including coordination strategies and impact on institutions and state higher education systems.

College Outcomes Evaluation Program Advisory Committee. (1987). *Report to the New Jersey Board of Higher Education from the Advisory Committee to the College Outcomes Evaluation Program.* Trenton, NJ: Department of Higher Education, Office of College Outcomes. The College Outcomes Evaluation Program (COEP) Advisory Committee was appointed in June 1985 by the New Jersey

Board of Higher Education for the purpose of developing a comprehensive higher education assessment program for New Jersey. The report suggests why a statewide assessment program should be developed in New Jersey and which areas of undergraduate education should be assessed, namely, basic skills, general intellectual skills, modes of inquiry, appreciation of the human condition/ethical issues, and the major. Emphasis is given to the need for institutional assessment of the personal development of students.

El-Khawas, E. (1991). *Campus trends, 1991. Higher Education Panel Report, No. 81.* Washington, D.C.: American Council on Education. A survey-research program established by the American Council on Education, the Higher Education Panel reports findings to policy makers in government, in the associations, and in educational institutions across the nation. The 1991 report, eighth in the series, gives special attention to the financial circumstances facing American higher education, particularly the nature, extent and early impact of budgetary cuts that have affected many colleges and universities.

Ewell, P.T. (1990). *Assessment and the "new accountability": A challenge for higher education's leadership.* Denver, CO: Education Commission of the States. A policy framework is proposed for understanding state-based assessment madates in terms of an emerging new conception of accountability. The framework is based on nine state case studies conducted in 1987-90 by the Education Commission of the States. Results are summarized in terms of an "external" and an "internal" agenda for action for higher education to make better use of assessment processes in building the public case for higher education.

Ewell, P.T., Finney, J.E., & Lenth, C. (1990). Filling in the mosaic: The merging pattern of state-based assessment. *AAHE Bulletin, 42,* 3-7. Summary results of a national survey of state-based assessment mandates and initiatives are reported. Detailed reviews of each state's effort are provided in a companion document by C. Poulsen (1990) published by the Education Commission of the States (Denver, CO).

National Governors' Association. (1988). *Results in education, state-level college assessment initiatives—1987-1988: Results of a 50-state survey.* Washington, D.C.: National Governors' Association. This report provides an overview of assessment in the fifty states based on responses to a survey conducted in the spring of 1988 for the purpose of determining what initiatives each state had undertaken to assess higher education outcomes.

Assessment in the Major

American Assembly of Collegiate Schools of Business. (1990). *Report of the AACSB Task Force on Outcome Measurement.* St. Louis, MO.: American Assembly of Collegiate Schools of Business. Sections cover historical context, goal analysis, and examples of goal statements. References are included also.

Banta, T.W., & Schneider, J.A, (1988). Using faculty-developed exit examinations to evaluate academic programs. *The Journal of Higher Education, 59,* 69-83. The experience of faculty at the University of Tennessee, Knoxville in developing examinations in the major field for purposes of assessing and improving curriculum and instruction is described. Difficulties encountered by the faculty, as well as the benefits they realized from the process, are identified and discussed.

Fong, B. (1988). Old wineskins: The AAC external examiner project. *Liberal Education, 74*, 12-16. The author describes a FIPSE-sponsored project involving eighteen institutions divided into six geographically proximate clusters of similar institutions. Faculty in three disciplines within each cluster cooperated in preparing and administering written and oral examinations to one anothers graduating seniors.

Stark, J.S., & Lowther, M.A. (1988). *Strengthening the ties that bind: Integrating undergraduate liberal and professional study.* Ann Arbor: The University of Michigan. This report of the professional preparation network identifies outcomes considered important by educators in eight undergraduate fields and suggests means of facilitating interdisciplinay communication.

Stone, H.L. and Meyer, T.C. (1989). *Developing an ability-based assessment program in the continuum of medical education.* Madison: University of Wisconsin Medical School. This is an instructional manual for developing programs which can go beyond measuring the acquisition of a knowledge base to identifying generic abilities that cross disciplinary lines but can be assessed with tasks unique to the disciplines. It contains twelve appendixes with sample criteria and rating forms.

Assessment in General Education

Alverno College Faculty (1979). *Assessment at Alverno College.* Milwaukee: Alverno Productions. Alverno College faculty have defined eight general student outcomes that form the basis of their instructional program and their assessment activities: effective commununications ability, analytical capability, problem-solving ability, valuing in decision-making, effective social interaction, effectiveness in individual/environmental relationships, responsible involvement in the contemporary world, and aesthetic responsiveness. Each of these outcomes has four levels at which students may attain certification through their coursework. This publication explains how the Alverno Faculty has designed and developed its assessment system.

Banta, T.W. (1991). Contemporary approaches to assessing general education outcomes. *Assessment Update, 3*(2), 1-4. The author argues for a planning-implementing assessing model for general education and illustrates each of these steps with current examples from institutional practice.

Forrest, A. (1986). *Good practices in general education.* Iowa City, Iowa: American College Testing Program. This loose-leaf collection provides examples from almost fifty institutions of good practice in stating goals for general education, setting curricular requirements, providing remedial support, orienting new students, designing courses, and writing examination questions.

Forrest, A. (1990). *Time will tell: Portfolio-assisted assessment of general education.* Washington, D.C.: American Association for Higher Education Assessment Fortun. This reference serves as a guide for colleges in establishing or improving the use of individual student portfolios in evaluating the general education program. The experience of seven institutions making extensive use of portfolios is related. Five sets of decisions form the basis for this work's organization: developing a definition of general education, deciding what to include in a portfolio and when, determining how and by whom the portfolios are to be analyzed, building a cost-effective process, and getting started.

Gaff, J.G., & Davis, M.L. (1981). Student views of general education. *Liberal Education, 67*(2), 112-113. A survey administered as part of the Project on General Education Models at twelve colleges and universities reveals student opinions about general education. The results indicate that students value a broad general education, especially as self-knowledge and preparation for a career. Students prefer some free choice, active learning methods, concreteness, and integration in their studies. They want to acquire communication competences, master thinking skills, and become proficient in personal and interpersonal relationships.

Classroom Assessment Techniques

Cross, K.P., & Angelo, T.A. (1988). *Classroom assessment techniques: A handbook for faculty.* Ann Arbor, Michigan: National Center for Research to Improve Postsecondary Teaching and Learning. The authors synthesized research literature concerning the impact of college on students in order to create a better focus for self-assessment of college teaching. The handbook describes thirty classroom techniques for assessing students' academic skills, learning skills, knowledge, self-awareness, and reactions to teaching and courses. Each technique comes with instructions for use, recommendations for analyzing data, suggestions for adaptation, and a list of pros and cons.

Developing Goals for Assessment

Gardiner, L.F. (1989). *Planning for assessment: Mission statements, goals, and objectives.* Trenton: New Jersey Department of Higher Education. This 255-page report is a step-by-step guide to planning for assessment, reviewing mission statements, and setting goals, with tables of sample outcomes goals and objectives. A chapter on resources includes reference materials, organizations, institutions with outcome-based programs, and an annotated bibliography.

Stark, J.S., Shaw, K.M., & Lowther, M.A. (1989). *Student goals for college and courses: A missing link in assessing and improving academic achievement. ASHE-ENC Higher Education Report, No. 6.* Washington, D,C.: Association for the Study of Higher Education. Getting students to take active responsibility for their own education may depend on whether or not what the students themselves hope to accomplish is taken into consideration. Helping students define and revise their own goals is a valid educational undertaking.

Assessment of Student Development

Hanson, G.R. (1989). *The assessment of student outcomes: A review and critique of assessment instruments.* Trenton, N.J.: New Jersey Department of Higher Education, College Outcomes Evaluation Program. This report, published to help faculty and administrators evaluate assessment instruments, provides an overview of assessment issues, strategies, instruments, and techniques. Hanson evaluates instruments that assess personal, career, moral, and cognitive development; student satisfaction; behavioral events; social attitudes; and learning styles.

Kozloff, J. (1987). A student-centered approach to accountability and assessment. *Journal of College Student Personnel. 28*(5), 419-424. Assessment of student outcomes has most often been described from the perspectives of faculty and administrators. The author maintains that the student perspective should also be considered. Procedures for developing a student-centered model are described.

Pace, C.R. (1990). *The undergraduates: A report of their activities and progress in college in the 1980s.* Los Angeles: University of California, Los Angeles, Center for the Study of Evaluation. Pace discusses five types of institutions, the six million undergraduates who attend them, the College Student Experiences Questionnaire that he developed to survey them, and the results of his survey of more than 25,000 students.

Assessment in Community Colleges

Bray, D., & Belcher, M.J. (1987). *Issues in student assessment: New Directions for Commununity Colleges, No. 59.* San Francisco: Jossey-Bass. This sourcebook contains 12 chapters by 15 different authors that deal with accountability issues and the political tensions they reflect; assessment practices, the use and misuse of testing, and emerging directions; and the impact of assessment, which includes issues of student access and opportunity, technological applications, expanded models for assessment, and increased linkages between high schools and colleges as a result of assessment information. The final challenge issued in the volume is to move assessment from political mandate to impetus for improving the curriculum and the quality of college teaching and learning.

Kreider, P.E., & Walleri, R.D. (1988). Seizing the agenda: Institutional effectiveness and student outcomes for community colleges. *Community College Review, 1*, 44-50. The authors urge community colleges to define their educational outcomes and use multi-dimensional approaches to assess their achievements. Examples of initiatives undertaken in California, Florida, and New Jersey are given.

League for Innovation in the Community College. (1990). *Assessing institutional effectiveness in community colleges.* Laguna Hills, California: League for Innovation in the Community College. The authors of this monograph encourage community colleges to build assessment programs around their stated missions. Chapters provide assessment guidelines for issues central to traditional community college missions, such as transfer, career preparation, basic skills development, continuing education, and access. Each mission is discussed in the context of clients, programs, assessing effectiveness, and assessing the mission itself. Two appendices discuss resources and instruments.

The National Alliance of Conununity and Technical Colleges. (1988). *Institutional effectiveness indicators.* This concise chart, showing major categories and specific indicators of institutional effectiveness in community and technical colleges, was developed by National Alliance members through an eighteen-month process.

Books/Collections/Review Articles on Assessment Topics

Astin, A.W. (1991). *Assessment for excellence.* New York: Macmillan Publishing Co. The author provides a detailed critique of traditional assessment policies and addresses the major issues related to assessment, including its underlying philosophy, classroom assessment, comprehensive assessment program design, statistical data analysis, and dissemination of results. A detailed explication of the input-environment-output (I-E-0) assessment model furnishes an organizing framework for the text.

Erwin, T.D. (1991). *Assessing student learning and development: A guide to the principles, goals, and methods of determining college outcomes.* San Francisco: Jossey-Bass Publishers, Inc. Intended as a place to start, this book is a practical guide for faculty, student affairs professionals, and administrators who are involved in assessing the effectiveness of their programs. The reference provides common terms, principles, issues, and selected literature from a variety of disciplines.

Ewell, P.T. (1988). Outcomes, assessment and academic improvement: In search of usable knowledge. In J.C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol IV, pp. 53-108). New York: Agathon Press. This chapter reviews the literature on assessment and undergraduate improvement, suggesting useful policy levers that can be employed by administrators at the institutional level. It contains a classification of issues and lessons from campus-based assessment programs reported in the literature through 1988.

Ewell, P.T. (1991). To capture the ineffable: New forms of assessment in higher education. In G. Grant (Ed.), *Review of Research in Education, 17,* 75-125. Washington, D.C.: American Educational Research Association. The author gives an historical and contextual account of assessment, concluding that an agenda of reform is in assessment's essential character. Ewell observes that assessment operates on two contradictory imperatives: academic improvement and external accountability. A taxonomy of approaches to assessment is offered, along with a critical review of current practice and recommendations for the future shape and direction of the field.

Farmer, D.W. (1988). *Enhancing student learning: Emphasizing essential competencies in academic programs.* Wilkes-Barre, Penn.: King's College Press. This monograph describes the implementation of an outcomes-oriented curriculum and course embedded assessment model at King's College. Peter Ewell remarks in his foreword to the book, "If the integrity of the curriculum is maintained and its effectiveness demonstrated, external benefits will naturally follow." A measure of that integrity at King's College is for faculty to be able to tell students how the intended outcomes of the curriculum are related to the college's definition of an educated person. Five years of faculty development preceded discussion of curriculum changes, with the concept of assessment introduced only after faculty had accepted an outcomesoriented curriculum. Excellence at King's College means measuring what actually happens to students while attending college and helping students transfer liberal learning skills across the curriculum.

Gray, P.J. (Ed.). (1989). Assessment goals and evaluation techniques. *New Directions for Higher Education, No. 67.* San Francisco: Jossey-Bass. This voluune focuses on analysis of the current state of assessment, differences between assessment and evaluation, guidelines, challenging aspects, and uses for information.

Halpern, D. (Ed.). (1987). Student outcomes assessment: What institutions stand to gain. *New Directions for Higher Education, No. 59.* San Francisco: Jossey-Bass. Halpern suggests in the overview that institutions change priorities in order to focus on student learning. Chapters include campus-based assessment programs, inadequacy of traditional measures, public policy issues, and models of student outcomes assessment

from Tennessee, California, Missouri, and New Jersey. In the final chapter Halpern notes eight factors that have been crucial to successful programs.

Hutchings, P., & Marchese, T. (1990). Watching assessment: Questions, stories, prospects. *Change, 22*, 12-38. The authors review the development of the assessment movement and consider how it is practiced. The assessment goals and programs of four campuses are described in detail as examples of how assessment is being carried out in diverse colleges and universities. Common characteristics, major themes, and recommendations for the future are offered.

Jacobi, M., Astin, A., & Ayala, F. (1987). *College student outcomes assessment: A talent development perspective*. Washington, D.C.: Association for the Study of Higher Education. This monograph describes several factors that conhibute to useful outcomes assessment: a) assessment data explicate issues facing educational practitioners, b) assessment yields information about students' growth and development, c) longitudinal data describe students' educational experiences so that their effects can be evaluated, and d) results are analyzed and presented in a manner that facilitates their use by practitioners. Approaches to assessment that make these contributions are described.

Light, R.J. (1990). *The Harvard Assessment Seminars: Explorations with students and faculty about teaching, learning, and student life. First Report.* Cambridge, MA: Harvard University, Graduate School of Education and Kermedy School of Government. The Harvard Assessment Seminars were begun in order to encourage innovation in teaching, curriculum, and advising, and to evaluate the effectiveness of each innovation. Administrators, faculty, and students work together in small groups on long-term research projects—designing, implementing, and assessing innovations. The report summarizes findings, lists participants and references, and provides descriptive charts.

Nichols, J.0. (1991). *A practitioners handbook for institutional effectiveness and student outcomes assessment implementation.* New York: Agathon Press. This book provides a general sequence of events for conducting outcomes assessment to evaluate institutional effectiveness. The model described by the author is intended to be adaptable to virtually all institutions and require only modest levels of funding. The book also contains detailed reviews of literature and practice related to critical points in the author's model.

Pace, C.R. (1979). *Measuring outcomes of college*. San Francisco: Jossey-Bass. Pace summarizes fifty years of research related to students' achievements during college and after college and the impact of college on student growth. He provides evidence that college graduates do develop in ways that differ from the development of nongraduates.

Pascarella, E.T., & Terenzini, P.T (1991). *How college affects students: Findings and insights from twenty years of research.* San Francisco: Jossey-Bass. Over 900 pages long, this volume contains nearly 200 pages of indexes and references, as well as a useful foreword and preface recommending ways to use the book. Introductory and summary chapters are provided for casual readers, while more invested readers may wish to read chapter summaries or entire chapters. Chapters cover issues such as theories and models of student change; development of attitudes, skills, morals, and values; psychosocial changes; educational attainment and career development; economic benefits and quality of life after college. A chapter on implications of the research for policy and practice concludes the book.

Seymour, D. (1991). *On Q: Causing quality in higher education*. New York: Macmillan. The author uses anecdotes from higher education and industry and quotations from interviews with chief academic officers from across the country to illustrate the principles of quality improvement. The book addresses such issues as defining quality, the differences between quality in industry and quality in higher education, the strategic implications of quality, communicating an emphasis on quality, the costs of quality, recruiting for quality, and the culture of quality in higher education.

Terenzini, P.T. (1989). Assessment with open eyes: Pitfalls in studying student outcomes. *Journal of Higher Education, 60,* 644-664. This article notes several purposes of assessment and analyzes issues such as involving administration and faculty, coordinating offices, determining political and practical effects, and calculating expenses. Also discussed are assets and limitations of different types of assessment measures and analyses of measures.

Wergin, J. F., & Braskamp, L.A. (Eds.). (1987). Evaluating administrative services and programs. *New Directions for Institutional Research, No. 56.* San Francisco: Jossey-Bass. This sourcebook suggests methods of assessing administrative and support programs as part of routine administrative practice.

Assessment Bibliographies

Center for Assessment Research and Development. (1990). *Bibliography of assessment instruments: Information on selected educational outcomes measures.* Knoxville: The University of Tennessee, Knoxville, Center for Assessment Research and Development. Assessment instrument reviews in this collection are grouped in the following categories: general education; basic skills; cognitive development; the major; values; and involvement, persistence, and satisfaction. Selected "Assessment Measures" columns from Assessment Update are also included.

Winthrop College Office of Assessment. (1990). *The SCHEA Network annotated bibliography of higher education assessment literature: Selected references from 1970-1989.* Winthrop, South Carolina: Winthop College Office of Assessment. The South Carolina Higher Education Assessment Network (SCHEA) collection presents works pertinent to issues of assessment of student development and learning in higher education. The articles are compiled under the following headings: assessment: general issues and principles; theoretical and educational aspects of assessment; assessment in the states; assessment of college readiness skills; assessment of general education; assessment of majors/specialty areas; assessment of career preparation; assessment of personal growth and development; data analysis and interpretation; survey methods; minority students: assessment and educational issues; non-traditional students; and retention and attrition.

Assessment Periodical

Assessment Update: Progress, trends, and practices in higher education. San Francisco: Jossey-Bass. This bi-monthly newsletter is edited at the Center for Assessment Research and Development at the University of Tennessee, Knoxville. The 16-page format of each issue includes several feature articles by leaders in the field of assessment, shorter articles on assessment methods being tried on individual campuses, notes on current publications, and colunms by Peter Ewell, Peter Gray and Gary Pike on state developments in assessment, campus assessment programs, and assessment instruments, respectively.

*__Trudy Banta__ is the Director of the Center for Assessment Research and Development at the University of Tennessee-Knoxville.

Council on Postsecondary Accreditation
One Dupont Circle, N.W., Suite 305
Washington, D.C. 20036
Telephone: 202/452-1433, Fax 202/331-9571